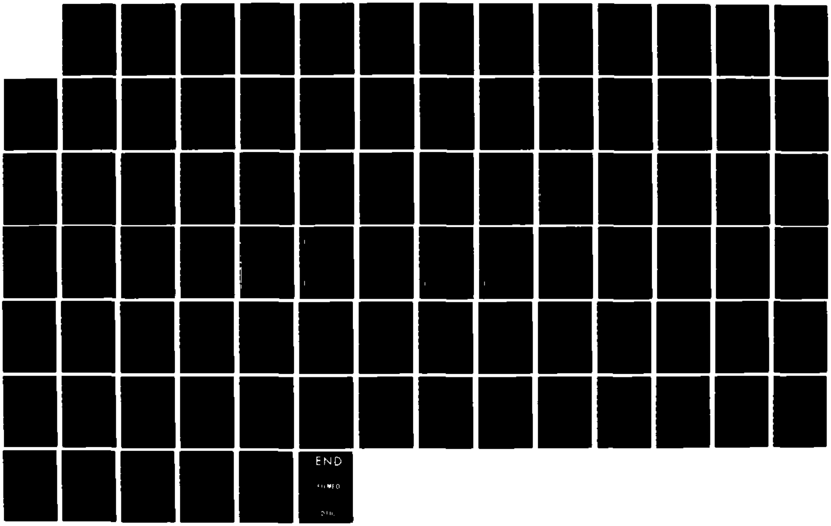
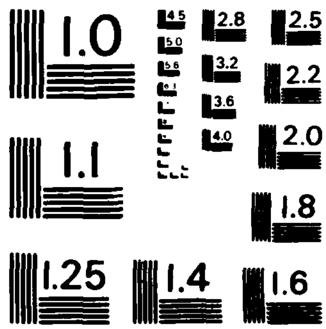
AD-A158 569 FURTHER INFORMATIONAL PROPERTIES OF THE NASH AND STACKELBERG SOLUTIONS OF..(U) UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES DEPT OF ELECTRI.. 1/1

UNCLASSIFIED G P PAPAVASSILOPOULOS 21 MAY 85 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963

(1') [full version of (1), submitted for publication] (under revision) (1)

Further Informational Properties of the Nash and Stackelberg Solutions of LQG Games

G.P.Papavassilopoulos and M.Tu
Dept.of Electrical Eng.-Systems
Univ.of Southern California
Los Angeles,CA 90089-0781

To DTIC
This is all the rept. as Final

Final Report

1 Introduction

In this paper we considers a two-decision-maker problem where each decision maker has his own information and ~~study~~ studies the impact of improving the information of only one decision maker.

In ~~[2]~~ a previous document an example of a two-decision-maker LQG static Nash game was considered and was shown for that particular example that, on the one hand, if one of the decision makers improves his own information by obtaining his opponent's information (while his opponent's information does not change) then he ends up with a higher Nash cost (Case B of [2]); on the other hand, if he improves his own information by getting an extra measurement not from his opponent (while his opponent's information does not change) then he might incur lower Nash cost (Case D of [2]). In this paper ~~we~~ proves that in a general two-decision-maker LQG static or dynamic Nash game, if one of the decision makers knows all his opponent's information, then more or better information for him alone is beneficial to him. In static games ~~we~~ the authors also prove that more

information for one of the decision makers alone is beneficial to him provided that such information is orthogonal to both decision makers' information. Additional keywords: numerical analysis; Kalman filtering; orthogonality; matrices (mathematics).

The structure of this paper is as follows: In Section 2 we study static games. By introducing the orthogonality condition of the information we give sufficient conditions that more information is beneficial to one of the decision makers. In Section 3 we formulate a two-decision-maker LQG dynamic Nash game where one of the decision maker's information is nested in the other's. At each stage k, decision maker 1 is allowed to use a function of estimates $\hat{x}_1(k)$ and $\hat{x}_3(k)$ of x(k) while decision maker 2 is allowed to use a function of $\hat{x}_1(k)$ only, where $\hat{x}_1(k)$ and $\hat{x}_3(k)$ are generated through two Kalman filters that use linear, noise-corrupted measurements of x(k) and $\hat{x}_3(k)$ is a refinement of $\hat{x}_1(k)$. In this setup the Nash solution exists, is unique and linear in $\hat{x}_1(k)$ and $\hat{x}_3(k)$ under certain invertibility assumptions on some matrices. Two nice features about the solution hold, namely, that a sort of separation principle of estimation and control holds and the estimation error is independent of the controls. In Section 4 we study the informational properties of the game formulated in Section 3. We prove that better information for decision maker 1 alone is beneficial to him. In Section 5 we extend the results obtained in Nash games to Stackelberg games. In Section 6 we give two examples to illustrate the informational properties discussed in the previous sections. Finally, in Section 7 we present our conclusions.

II. Some Informational Properties of LQG Static Nash Games

Consider a two-decision-maker LQG static Nash game. The cost functional of decision maker i, i=1,2 is denoted by†

$$J_i(\gamma_1, \gamma_2) = E[x'P_i'u_i + \frac{1}{2} u_i'u_i + u_i'Q_i u_j] \tag{1}$$

$$j \neq i, \quad i, j = 1, 2$$

where $x \in R^n$ is a Gaussian random vector, $x \sim N(0, \Omega)$, $u_i \in R^{\ell_i}$ is the control variable of decision maker i and P_i, Q_i are real constant matrices of appropriate dimensions. The linear measurement of decision maker i is given by

$$y_i = H_i x + \omega_i \tag{2}$$

H_i is an $m_i \times n$ real constant matrix and ω_i is a Gaussian random vector, $\omega_i \sim N(0, \Sigma_i)$ which is independent of x and of ω_j $(j \neq i)$. The control law γ_i is chosen from Γ_i where Γ_i consists of all the measurable functions from R^{m_i} to R^{ℓ_i} such that $\gamma_i(y_i)$ is a second order random vector. A pair (γ_1^*, γ_2^*) is called a Nash solution of the game if it satisfies the following two inequalities

$$J_1(\gamma_1^*, \gamma_2^*) \leq J_1(\gamma_1, \gamma_2^*) \tag{3a}$$

$$J_2(\gamma_1^*, \gamma_2^*) \leq J_2(\gamma_1^*, \gamma_2) \tag{3b}$$

for every $\gamma_1 \in \Gamma_1$ and $\gamma_2 \in \Gamma_2$. γ_i^* is called the Nash strategy of decision maker i. A necessary and sufficient condition characterizing a Nash solution of the above game was given in Theorem 1 of [3] which we state below as a lemma.

Lemma 2.1. A pair (γ_1^*, γ_2^*) is a Nash solution of the game described above if and only if the following two equalities hold

$$\gamma_i^*(y_i) = -P_i E[x \mid y_i] - Q_i E[\gamma_j^*(y_j) \mid y_i] \qquad (4)$$
$$j \neq i, \quad i, j = 1, 2$$

Using Lemma 2.1 we will show how the Nash solution is affected by the information available to the decision makers and hence how the Nash performance is affected by the information structure. We need the following definition of orthogonality and a lemma which consists of several well-known facts in estimation theory [4].

Definition 2.2. Two zero-mean Gaussian random vectors z_1 and z_2 are said to be orthogonal (denoted by $z_1 \perp z_2$) if $E[z_1 z_2'] = 0$. Two sets Z_1 and Z_2 are orthogonal if $z_1 \perp z_2$ for every $z_1 \in Z_1$ and $z_2 \in Z_2$.

Lemma 2.3. Let z_i, i=1,2,3 be zero-mean Gaussian random vectors, then

(i) $\{z_1 - E[z_1 \mid z_2]\} \perp z_2$

(ii) $E[z_1 \mid z_2] = Cz_2$ where C is a real matrix.

If, in addition, $z_2 \perp z_3$, then

(iii) $E[z_2 \mid z_3] = 0$

(iv) $E[z_1 | z_2, z_3] = E[z_1 | z_2] + E[z_1 | z_3]$.

Denote an extra measurement by y_e,

$$y_e = H_e x + \omega_e \tag{5}$$

where H_e is an $m_e \times n$ real constant matrix and ω_e is a Gaussian random vector, $\omega_e \sim N(0, \Sigma_e)$ and is independent of x, w_1, w_2.

Condition C (i) $y_e \perp \{y_1, y_2\}$, (ii) $y_2 = My_1$

where M is an $m_2 \times m_1$ matrix. The meaning of Condition C (ii) is that the information provided by y_2 is contained in that provided by y_1.

Lemma 2.4. Under either one of Conditions C,
$\{E[x | y_1, y_e] - E[x | y_1]\} \perp \{y_1, y_2\}$.

Proof: Under Condition C (i), Lemma 2.3 (iv) and (ii) imply that

$$E[x | y_1, y_e] - E[x | y_1] = E[x | y_1] + E[x | y_e] - E[x | y_1]$$
$$= E[x | y_e] = Cy_e \tag{6}$$

The result holds since $y_e \perp \{y_1, y_2\}$.

Under Condition C (ii),

$$E[x | y_1, y_e] - E[x | y_1] = E[x | y_1, y_e] - E[E[x | y_1, y_e] | y_1] \tag{7}$$

and thus (Lemma 2.3 (i)) $\{E[x | y_1, y_e] - E[x | y_1]\} \perp y_1$ and by Condition C (ii) $\{E[x | y_1, y_e] - E[x | y_1]\} \perp y_2$. $\square$

The question of existence and uniqueness of the Nash solution has been studied in [3] and [5] where it was shown that almost always there exists a unique solution which has to be an affine function of the information. more offici...

Theorem 2.5. Let at least one of Condition C (i), (ii) hold, then if there exists a Nash solution under the information pattern where decision maker 1 knows y_1 and decision maker 2 knows y_2 then there exists a Nash solution under the information pattern where decision maker 1 knows (y_1, y_e) and decision maker 2 knows y_2 and vice versa. Furthermore, the Nash strategy γ_2 is the same under both information patterns. For the case where Condition C (ii) holds, a Nash solution exists and is unique if and only if the matrix $I-Q_1Q_2$ is invertible.

Proof: (i) Let Condition C (i) hold.

When decision maker 1 knows y_1 and decision maker 2 knows y_2, by Lemma 2.1, a Nash solution $(\gamma_1(y_1), \gamma_2(y_2))$ exists if and only if

$$\gamma_2(y_2) = Q_2Q_1E[E[\gamma_2(y_2)\,|\,y_1]|y_2]+Q_2P_1E[E[x|y_1]|y_2]-P_2E[x|y_2] \tag{8}$$

When decision maker 1 knows $\{y_1, y_e\}$ and decision maker 2 knows y_2, a Nash solution $(\gamma_1(y_1, y_e), \gamma_2(y_2))$ exists if and only if

$$\gamma_2(y_2) = Q_2Q_1E[E[\gamma_2(y_2)\,|y_1, y_e]|y_2]+Q_2P_1E[E[x|y_1,y_e]|y_2]-P_2E[x|y_2] \tag{9a}$$

$$= Q_2Q_1E[E[\gamma_2(y_2)\,|y_1]|y_2]+Q_2P_1E[E[x|y_1]|y_2]-P_2E[x|y_2] \tag{9b}$$

where we use Lemma 2.3 (iii) and (iv) and the fact that $\gamma_2(y_2)$ is affine in y_2. Equations (8) and (9b) are exactly the same hence we have the desired result.

(ii) Let Condition C (ii) hold.

When decision maker 1 knows y_1 and decision maker 2 knows y_2, a Nash solution $(\gamma_1(y_1), \gamma_2(y_2))$ exists if and only if

$$\gamma_2(y_2) = Q_2Q_1\gamma_2(y_2) + Q_2P_1E[x|y_2] - P_2E[x|y_2] \tag{10}$$

When decision maker 1 knows $\{y_1, y_e\}$ and decision maker 2 knows y_2, a Nash solution $(\gamma_1(y_1, y_e), \gamma_2(y_2))$ exists if and only if

$$\gamma_2(y_2) = Q_2Q_1\gamma_2(y_2) + Q_2P_1E[x|y_2] - P_2E[x|y_2] \tag{11}$$

Equations (10) and (11) are the same and hence if a Nash solution exists in one of the information patterns, it exists in the other and γ_2 is the same in both information patterns. Furthermore, a unique Nash solution exists if and only if $I - Q_1Q_2$ is invertible. □

Theorem 2.6. Let Condition C (i) or (ii) hold, then the Nash cost incurred to decision maker 1 when the information available to decision maker 1 is $\{y_1, y_e\}$ and to decision maker 2 is y_2 is less than or equal to the Nash cost incurred to decision maker 1 when the information available to decision maker 1 is y_1 and to decision maker 2 is y_2.

Proof: Let (γ_1^*, γ_2^*) denote the Nash solution when decision maker 1 knows $\{y_1, y_e\}$ and decision maker 2 knows y_2 and (γ_1^0, γ_2^0) the Nash solution when decision maker 1 knows y_1 and decision maker 2 knows

y_2, then by Theorem 2.5

$$\begin{aligned} J_1(\gamma_1^*, \gamma_2^*) &= J_1(\gamma_1^*, \gamma_2^0) \\ &= \min_{\gamma_1(y_1, y_e)\in\Gamma_1'} J_1(\gamma_1, \gamma_2^0) \\ &\leq \min_{\gamma_1(y_1)\in\Gamma_1} J_1(\gamma_1, \gamma_2^0) \\ &= J_1(\gamma_1^0, \gamma_2^0) \end{aligned} \tag{12}$$

where Γ_1' consists of all the measurable functions from $R^{m_1+m_e}$ to R^{ℓ_1}.

□

<u>Remark 2.7.</u> Notice that Theorem 2.5 and 2.6 hold regardless of the functional form of the costs as long as they are quadratic.

<u>Remark 2.8.</u> All the results obtained in this section go through even if we assume that x is not of zero mean. This is easy to verify.

III. <u>Formulation of an LQG Dynamic Nash Game and Its Solution</u>

Consider a two-decision-maker, N-stage Nash game where the state of the system $x(\cdot)$ evolves according to

$$x(k+1) = Ax(k) + B_1u_1(k) + B_2u_2(k) + \omega(k), \quad x(0) = x_0 \tag{17}$$

where $k\in\theta_i = \{0, 1, \ldots, N-1\}$, $x(k)\in R^n$ and $u_i(k)\in R^{\ell_i}$ denotes the control variable of decision maker i at stage k, i=1, 2. x_0 and $\{\omega(k), k\in\theta\}$ are independent Gaussian random vectors, $x_0 \sim N(\bar{x}_0, \Omega_0)$, $\omega(\cdot) \sim N(0, R)$.

At each stage k, the~~re are~~ measurements $y_i(k) \in R^{m_i}$, $i=1,2$, are given by

$$y_i(k) = H_i x(k) + v_i(k) \tag{18}$$

where $\{v_i(k),\ k \in \theta_1,\ i=1,2\}$ are independent Gaussian random vectors, $v_i(\cdot) \sim N(0, \Sigma_i)$. v_i's are also independent of x_0 and $\{\omega(k),\ k \in \theta_1\}$. The information available to the decision makers is not $y_i(k)$'s, but $\hat{x}_1(k)$, $\hat{x}_3(k)$, the estimates of $x(k)$ given by two Kalman filters:

$$\hat{x}_i(k) = \hat{x}_i(k/k-1) + G_i(k)[y_i(k) - H_i\hat{x}_i(k/k-1)] \tag{19a}$$

$$\hat{x}_i(k+1/k) = A\hat{x}_i(k) + B_1u_1(k) + B_2u_2(k), \quad \hat{x}_i(0/-1) = \bar{x}_0 \tag{19b}$$

$$G_i(k) = \Sigma_i(k/k-1)H_i'(H_i\Sigma_i(k/k-1)H_i' + \Sigma_i)^{-1} \tag{19c}$$

$$\Sigma_i(k+1/k) = A[I-G_i(k)H_i]\Sigma_i(k/k-1)A' + R, \quad \Sigma_i(0/-1) = \Omega_0 \tag{19d}$$

$$\Sigma_i(k) = [I - G_i(k)H_i]\Sigma_i(k/k-1) \tag{19e}$$

$i = 1, 3$

where

$$H_3 \triangleq [H_1', H_2']' \tag{20a}$$

$$y_3(\cdot) \triangleq [y_1'(\cdot), y_2'(\cdot)]' \tag{20b}$$

$$\Sigma_3 \triangleq \mathrm{diag}\,[\Sigma_1, \Sigma_2] \tag{20c}$$

$\hat{x}_i(k+1/k)$ is the one-step prediction estimate and $\Sigma_i(k)$ and $\Sigma_i(k+1/k)$ are the error covariance matrices associated with $\hat{x}_i(k)$ and $\hat{x}_i(k+1)$, respectively,

$$\Sigma_i(k) = E\{[x(k) - \hat{x}_i(k)][x(k) - \hat{x}_i(k)]'\} \tag{21}$$

$$\Sigma_i(k+1/k) = E\{[x(k+1) - \hat{x}_i(k+1/k)][x(k+1) - \hat{x}_i(k+1/k)]'\} \tag{22}$$

The information structure is defined as follows: At each stage k, decision maker 1 knows $I_1(k) \triangleq \{\hat{x}_1(k), \hat{x}_3(k)\}$ while decision maker 2 knows $I_2(k) \triangleq \{\hat{x}_1(k)\}$. This information structure can be justified by considering that there are two impartial referees 1 and 3 who compute respectively $\hat{x}_1(k)$ and $\hat{x}_3(k)$, referee 1 gives $\hat{x}_1(k)$ to both decision makers and referee 3 gives $\hat{x}_3(k)$ to decision maker 1 only.

The cost of decision maker i is J_i ($\triangleq J_i(0)$) where $J_i(k)$ denotes the cost to go of decision maker i at stage k and is defined by

$$J_i(k) = E\left\{\sum_{n=k}^{N-1}[x'(n)P_i x(n) + u_i'(n)u_i(n) + u_j'(n)Q_i u_j(n)] + x'(N)P_i x(N)\right\} \tag{23}$$
$$j \neq i, \ i,j = 1,2$$

where $P_i, Q_i \geq 0$. $u_i(k)$ is chosen as $\gamma_i^k(I_i(k))$ and the γ_i^k's are measurable functions, $\gamma_1^k: R^n \times R^n \to R^{\ell_1}$ and $\gamma_2^k: R^n \to R^{\ell_2}$ with the property that $\gamma_i^k(I_i(k))$ is a second order random vector.

Let

$$g_i \triangleq \{\gamma_i^0, \gamma_i^1, \ldots, \gamma_i^{N-1}\}, \quad i = 1,2 \tag{24}$$

A pair $\{g_1^*, g_2^*\}$ is called a Nash solution of the game if

$$J_1(g_1^*, g_2^*) \leq J_1(g_1, g_2^*) \qquad \forall \text{ admissible } g_1 \tag{25a}$$

$$J_2(g_1^*, g_2^*) \leq J_2(g_1^*, g_2) \qquad \forall \text{ admissible } g_2 \tag{25b}$$

Before we give the Nash solution of the game, we need the following lemma which shows an orthogonality in the information structure and the proof is given in Appendix A.

Lemma 3.1 (i) $E[\hat{x}_3(k) \mid \hat{x}_1(k)] = \hat{x}_1(k)$.

Let $\tilde{\hat{x}}_4(k) = \hat{x}_3(k) - \hat{x}_1(k)$ then

(ii) $\hat{x}_4(k) \perp \hat{x}_1(k)$ and

(iii) $E[\hat{x}_4(k)] = 0$, $E[\hat{x}_4(k), \hat{x}_4'(k)] = \Sigma_1(k) - \Sigma_3(k)$.

Notice that by Lemma 3.1, the information structure $I_1(k)$ can equivalently be considered as $I_1(k) = \{\hat{x}_1(k), \hat{x}_4(k)\}$ which consists of two orthogonal elements.

The Nash solution of the game described above is provided in the following theorem, the proof of which is given in Appendix B.

Theorem 3.2. Consider the equations

$$\begin{aligned} L_i(k) = {} & P_i + A'[(I+B_1B_1'L_1(k+1)+B_2B_2'L_2(k+1))^{-1}]'[L_i(k+1)+L_i(k+1)B_iB_i'L_i(k+1) \\ & + L_j(k+1)B_jQ_i B_j' L_j(k+1)][I+B_1B_1'L_1(k+1)+B_2B_2'L_2(k+1)]^{-1}A, \\ & L_i(N) = P_i, \quad j \neq i, \quad i,j = 1,2 . \end{aligned} \tag{26}$$

which evolve backwards in time. We assume the inverse of $(I+B_1B_1'L_1(k) + B_2B_2'L_2(k))$ exists for every $k \in \theta$, then

(i) There exists a unique Nash solution to the game which is the following:

$$u_1^*(k) = {}^*\gamma_1^k(I_1(k)) = F_{11}(k)\hat{x}_1(k) + F_{14}(k)\hat{x}_4(k) \tag{27}$$

$$u_2^*(k) = {}^*\gamma_2^k(I_2(k)) = F_2(k)\hat{x}_1(k) \tag{28}$$

where

$$F_{11}(k) = -B_1'L_1(k+1)[I+B_1B_1'L_1(k+1) + B_2B_2'L_2(k+1)]^{-1}A \tag{29}$$

$$F_{14}(k) = -B_1'L_1(k+1)[I+B_1B_1'L_1(k+1)]^{-1}A \tag{30}$$

$$F_2(k) = -B_2'L_2(k+1)[I+B_1B_1'L_1(k+1)+B_2B_2'L_2(k+1)]^{-1}A \tag{31}$$

(ii) The cost to go of decision maker i at stage k is

$$J_i(k) = E[\hat{x}_3'(k)L_i(k)\hat{x}_3(k)] + K_i(k) \tag{32}$$

where

$$K_i(k) = \mathrm{tr}\Big\{[A'L_i(k+1)A - L_{i4}(k)+P_i]\Sigma_3(k) - L_i(k+1)[\Sigma_3(k+1)-R] + L_{i4}(k)\Sigma_1(k)\Big\}$$
$$+ K_i(k+1)\ , \quad K_i(N) = \mathrm{tr}\Big\{P_i\Sigma_3(N)\Big\} \tag{33}$$

$$L_{14}(k) = A'[(I+B_1B_1'L_1(k+1))^{-1}]'[L_1(k+1)+L_1(k+1)B_1B_1'L_1(k+1)]$$
$$[I+B_1B_1'L_1(k+1)]^{-1}A - L_1(k) + P_1 \tag{34}$$

$$L_{24}(k) = A'[(I+B_1B_1'L_1(k+1))^{-1}]'[L_2(k+1)+L_1(k+1)B_1Q_2B_1'L_1(k+1)]$$
$$[I+B_1B_1'L_1(k+1)]^{-1}A - L_2(k) + P_2 \tag{35}$$

<u>Remark 3.3.</u> Notice that the control laws $F_{11}(k)$, $F_{14}(k)$ and $F_2(k)$ in the above theorem are independent of the observation noise

in the measurements (18), i.e., a sort of separation principle holds under such information structure. Also we can see from the Kalman filter equations (19) that the estimation error $\Sigma_i(k)$, $i=1,3$ is independent of the controls.

Remark 3.4. Compare now Theorem 3.2 (where $I_1(k) = \{\hat{x}_1(k), \hat{x}_3(k)\}$, $I_2(k) = \{\hat{x}_1(k)\}$) with Theorem 2.1 of [1] (where $I_1(k) = \{\hat{x}_1(k)\}$, $I_2(k) = \{\hat{x}_1(k)\}$ and let $M=2$) we see that ${}^*\gamma_2^k(I_2(k))$, the Nash strategy of decision maker 2 is the same in the two different information structures. (~~In case $\hat{x}_1(k) = \hat{x}_3(k)$ (i.e., $\hat{x}_4(k) = 0$) $\forall k \in \theta_1$, then Theorem 4.3.2 is exactly the same as Theorem 3.2.1 with $M=2$.~~) Motivated by Theorem 2.5 and 2.6 where we see that more information to the decision maker who knows all his opponent's information is beneficial to him, we expect that the extra information $\hat{x}_4(k)$ ($I_1(k)$ compared with $I_2(k)$ in Theorem 3.2) is beneficial to decision maker 1, which is indeed true and will be shown in the following section.

Remark 3.5. The nonsingularity condition of the matrix $I = B_1B_1'L_1(k) + B_2B_2'L_2(k)$ and the boundedness condition of $L_i(k)$, the solution of the coupled Riccati equations (31) were discussed in Theorem 2.2 and Remark 2 of [1].

IV Some Informational Properties of LQG Dynamic Nash Games

In this section we first give the definition of "better information for decision maker 1 alone," then compare the Nash costs of both

decision makers resulting from two different information and then prove that better information for decision maker 1 alone is beneficial to him. A sufficient condition that better information for decision maker 1 alone is beneficial to decision maker 2 is also derived.

Consider Information I and II. In Information I the estimates $\hat{x}_1^I(k)$ and $\hat{x}_3^I(k)$ are generated through the past controls and the measurements

$$y^I(\cdot) = H_i^I x(\cdot) + v_i^I(\cdot) , \qquad v_i^I \sim N(0, \Sigma_i^I) \tag{36}$$

$$i = 1, 2 ,$$

with corresponding estimation error $\Sigma_1^I(k)$ and $\Sigma_3^I(k)$. In Information II the estimates $\hat{x}_1^{II}(k)$ and $\hat{x}_3^{II}(k)$ are generated through the past controls and the measurements

$$y_i^{II}(\cdot) = H_i^{II} x(\cdot) + v_i^{II}(\cdot) , \qquad v_i^{II} \sim N(0, \Sigma_i^{II}) \tag{37}$$

$$i = 1, 2 ,$$

with corresponding estimation error $\Sigma_1^{II}(k)$ and $\Sigma_3^{II}(k)$.

Definition 4.1. We say that Information I provides better information for decision maker 1 alone than Information II if $\Sigma_1^I(k) = \Sigma_1^{II}(k)$, $\Sigma_3^I(k) \leqq \Sigma_3^{II}(k)$ for every $k \in \theta_1$ and $\Sigma_3^I(k) \neq \Sigma_3^{II}(k)$ for at least one $k \in \theta_1$.

An obvious fact about the definition given above is that all the improvement is in the part of $\hat{x}_4(\cdot)$, decision maker 1's private information while there is no improvement in the part of $\hat{x}_1(\cdot)$, the public information of both decision makers.

Let $J_i^I(k)$ and $K_i^I(k)$, i=1, 2, be defined as in (32) and (33)

corresponding to Information I and

$$\hat{\Omega}_j^I(k) = E[\hat{x}_j^I(k)\,\hat{x}_j^{I'}(k)]\,, \quad j = 1, 3\,.$$

Similarly we define $J_i^{II}(k)$, $K_i^{II}(k)$, $i=1,2$ and $\hat{\Omega}_j^{II}(k)$, $j = 1, 3$ for Information II.

Theorem 4.3 The Nash solution given by Theorem 4.3.2 has the property that better information for decision maker 1 alone does not increase decision maker i's cost if

$$P_i + A' L_i(k+1) A - L_i(k) - L_{i4}(k) \geqq 0 \qquad \text{for every } k \in \theta_1\,. \tag{38}$$

It lowers decision maker i's cost with strict inequality in (38).

Proof: From part (ii) of Theorem 3.2,

$$J_i^I(0) = E[\hat{x}_3^{I'}(0) L_i(0) \hat{x}_3^I(0)] + K_i^I(0) = \operatorname{tr}\{L_i(0)\,\hat{\Omega}_3^I(0)\} + K_i^I(0)\,. \tag{39}$$

From the recursive expression of $K_i(\cdot)$ in (33) we obtain

$$J_i^I(0) = \operatorname{tr}\Big\{L_i(0)\hat{\Omega}_3^I(0) + [P_i + A'L_i(1)A - L_{i4}(0)]\Sigma_3^I(0) + L_i(1)R + L_{i4}(0)\Sigma_1^I(0) + \sum_{k=1}^{N-1}\Big[[P_i + A'L_i(k+1)A - L_i(k) - L_{i4}(k)]\Sigma_3^I(k) + L_i(k+1)R + L_{i4}(k)\Sigma_1^I(k)\Big]\Big\} \tag{40}$$

Similarly

$$J_i^{II}(0) = \mathrm{tr}\Big\{L_i(0)\hat{\Omega}_3^{II}(0)+[P_i+A'L_i(1)A-L_{i4}(0)]\Sigma_3^{II}(0)+L_i(1)R+L_{i4}(0)\Sigma_1^{I}(0)+$$

$$\sum_{k=1}^{N-1}\Big[[P_i+A'L_i(k+1)A-L_i(k)-L_{i4}(k)]\Sigma_3^{I}(k)+L_i(k+1)R+L_{i4}(k)\Sigma_1^{I}(k)\Big]\Big\} \qquad (41)$$

By using the fact that

$$\hat{\Omega}_3^{II}(0) - \hat{\Omega}_3^{I}(0) = -(\Sigma_3^{II}(0) - \Sigma_3^{I}(0)) \qquad (42)$$

we obtain

$$J_i^{II}(0) - J_i^{I}(0) = \sum_{k=0}^{N-1} \mathrm{tr}\Big\{[P_i+A'L_i(k+1)A-L_i(k)-L_{i4}(k)][\Sigma_3^{II}(k)-\Sigma_3^{I}(k)]$$

$$+ L_{i4}(k)[\Sigma_1^{II}(k) - \Sigma_1^{I}(k)]\Big\} \qquad (43)$$

Suppose now that Information I provides better information for decision maker 1 alone than Information II, then Lemma 4.2 implies $J_i^{II}(0) \geqq J_i^{I}(0)$ if

$$P_i + A'L_i(k+1)A - L_i(k) - L_{i4}(k) \geqq 0 \quad \text{for every } k\in\theta_1 , \qquad (44)$$

and $J_i^{II}(0) > J_i^{I}(0)$ if the inequality is strict in (44).

Corollary 4.4. Better information for decision maker 1 alone does not increase decision maker 1's Nash cost. It lowers decision maker 1's Nash cost provided that the matrices A, B_1B_1' and P_1 are nonsingular.

Proof: Substituting (26) and (34) into (44), we obtain

$$P_1 + A'L_1(k+1)A - L_1(k) - L_{14}(k)$$

$$= A'L_1(k+1)A - A'[(I+B_1B_1'L_1(k+1))^{-1}]'[L_1(k+1)+L_1(k+1)B_1B_1'L_1(k+1)]$$
$$[I+B_1B_1'L_1(k+1)]^{-1}A$$

$$= U'VU \geq 0 \tag{45}$$

where

$$U = [I + B_1B_1'L_1(k+1)]^{-1}A \tag{46}$$

and

$$V = L_1(k+1)B_1B_1'L_1(k+1) + L_1(k+1)B_1B_1'L_1(k+1)B_1B_1'L_1(k+1) \geq 0 \tag{47}$$

Furthermore, if $P_1 > 0$ then (26) implies that $L_1(k) > 0$, hence U is nonsingular and $V > 0$ provided that A, B_1B_1' and P_1 are nonsingular. Theorem 4. 3. then implies the desired result.

Remark 4. 5. Notice the resemblance of equation (47) to ~~(38)~~ of [1]. This is so since $\hat{x}_4(\cdot)$ is orthogonal to decision maker 2's information, any improvement in the part of $\hat{x}_4(\cdot)$ is totally used by decision maker 1 to optimize his performance which brings forth the team-like benefit.

Remark 4.6 In Corollary 4.4 we see that better information for decision maker 1 alone is beneficial to him and this fact is independent of the number of stages N and it is not necessary for the "better" information to be "dynamically better." In contrast with Theorem 5.19[1] the above two features reveal the essential

difference between improving the decision makers' "private" information and "public" information in a dynamic Nash game.

V. Related Properties of Static and Feedback Stackelberg Games

In this section we extend the results obtained in Nash games to static and feedback Stackelberg games. The difference of a Stackelberg game and a Nash game lies partially in that the roles of the decision makers are asymmetric in Stackelberg games while it is symmetric in a Nash game. However, the Stackelberg solution of a static game is also a Nash solution of the same problem under explicit control sharing and a feedback Stackelberg solution of an N-stage dynamic game is also a Nash solution of a 2N-stage game (as has been observed in [6]). Hence we expect some different as well as some similar properties ~~in~~ between Stackelberg ~~games as in~~ and Nash games.

Consider a two-decision-maker static Stackelberg game. Let decision maker 1 be the leader and decision maker 2 the follower. Their cost functionals are given by $J_1(\gamma_1, \gamma_2)$ and $J_2(\gamma_1, \gamma_2)$, respectively, where

$$J_i(\gamma_1, \gamma_2) = E\left[\frac{1}{2}u_i'u_i + \frac{1}{2}u_j'P_iu_j + u_i'Q_iu_j + u_i'S_{ii}x + u_j'S_{ij}x\right] \tag{48}$$

$$j \neq i, \quad i, j = 1, 2 .$$

where $x \in R^n$ is a Gaussian random vector, $x \sim N(0, \Omega)$, $u_i \in R^{\ell_i}$ is the control variable of decision maker i and P_i, Q_i, S_{ii} and S_{ij} are real constant matrices of appropriate dimensions. The linear

measurement of decision maker i is given by

$$y_i = H_i x + \omega_i \quad . \tag{49}$$

H_i is an $m_i \times n$ real constant matrix and ω_i is a Gaussian random vector, $\omega_i \sim N(0, \Sigma_i)$ which is independent of x. The control law γ_i is chosen from Γ_i where Γ_i consists of all the measurable functions mapping from R^{m_i} to R^{ℓ_i} such that $\gamma_i(y_i)$ is a second order random vector. A pair (γ_1^*, γ_2^*) is called a Stackelberg solution with decision maker 1 as the leader if γ_1^* satisfies the following inequality

$$\sup_{\gamma_2 \in R_2(\gamma_1^*)} J_1(\gamma_1^*, \gamma_2) \leqq \sup_{\gamma_2 \in R_2(\gamma_1)} J_1(\gamma_1, \gamma_2) \tag{50}$$

for every $\gamma_1 \in \Gamma_1$ and $\gamma_2^* \in R_2(\gamma_1^*)$, where $R_2(\gamma_1)$ is called the rational reaction set of the follower to the strategy γ_1 announced by the leader, and it is defined by

$$R_2(\gamma_1) = \left\{ \gamma_2^0 \in \Gamma_2 \,\middle|\, J_2(\gamma_1, \gamma_2^0) \leqq J_2(\gamma_1, \gamma_2) \ , \ \forall \gamma_2 \in \Gamma_2 \right\} \tag{51}$$

Notice that if $R_2(\gamma_1)$ is a singleton for each $\gamma_1 \in \Gamma_1$, then (50) can equivalently be written as

$$J_1(\gamma_1^*, \gamma_2^0(\gamma_1^*)) \leqq J_1(\gamma_1, \gamma_2^0(\gamma_1)) \quad . \tag{52}$$

It turns out that $R_2(\gamma_1)$ is a singleton indeed [7] and is given by

$$\gamma_2^0(\gamma_1, y_2) = -S_{22}E[x \mid y_2] - Q_2 E[\gamma_1(y_1) \mid y_2] \quad . \tag{53}$$

A sufficient condition that a unique linear Stackelberg solution exists was given in [7] which condition is determined by the matrices P_i and Q_i, $i=1,2$, and has nothing to do with the information available to the decision makers. We assume, in the following derivations that a unique linear Stackelberg solution exists under every information we will consider. The result of the following lemma is known but we include a short proof for reasons of completeness.

Lemma 5.1. The leader's cost decreases if he has an extra measurement y_e available.

Proof: Let (γ_1^*, γ_2^*) and (γ_1^0, γ_2^0) denote respectively the Stackelberg solution before and after the leader acquires y_e. After the leader acquires y_e, he can choose a suboptimal strategy $\gamma_1^S(y_1, y_e) = \gamma_1^*(y_1)$, then the follower will react by choosing $\gamma_2^S(y_2) = \gamma_2^*(y_2)$ and hence

$$J_1(\gamma_1^0(y_1, y_e), \gamma_2^0(y_2) \leqq J_1(\gamma_1^S(y_1, y_e), \gamma_2^S(y_2)$$

$$= J_1(\gamma_1^*(y_1), \gamma_2^*(y_2)) \tag{54}$$

□

The follower, who is in the lower level of a hierarchy, see things different from the leader and knowing more is not necessarily beneficial to him. As in the Nash case, we first prove in the following theorem that if the follower acquires extra measurement y_e which satisfies certain orthogonality conditions or the follower knows all that the leader knows, then such y_e is beneficial to the follower.

Condition $\bar{C}$ (i) $y_e \perp \{y_1, y_2\}$, (ii) $y_1 = My_2$

Theorem 5.2. If the follower acquires extra measurement y_e such that either one of Condition $\bar{C}$ holds, then the leader's strategy does not change.

Proof: Let $\gamma_1^*(y_1)$, $\gamma_1^0(y_1)$ denote the leader's strategy before and after the follower acquires y_e and $\gamma_2^*(\gamma_1, y_2)$, $\gamma_2^0(\gamma_1, y_2, y_e)$ denote respectively the follower's reaction before and after he acquires y_e, then by (53)

$$\gamma_2^*(\gamma_1, y_2) = -S_{22}E[x|y_2] - Q_2E[\gamma_1(y_1)|y_2], \tag{55}$$

and

$$\gamma_2^0(\gamma_1, y_2, y_e) = -S_{22}E[x|y_2, y_e] - Q_2E[\gamma_1(y_1)|y_2, y_e]. \tag{56}$$

Under either one of Conditions $\bar{C}$ the following is true

$$E[\gamma_1(y_1)|y_2, y_e] = E[\gamma_1(y_1)|y_2]. \tag{57}$$

Hence (56) can be written as

$$\begin{aligned}\gamma_2^0(\gamma_1, y_2, y_e) &= \gamma_2^*(\gamma_1, y_2) - S_{22}\{E[x|y_2, y_e] - E[x|y_2]\} \\ &= \gamma_2^*(\gamma_1, y_2) - S_{22}\hat{y}\end{aligned} \tag{58}$$

where $\hat{y} \triangleq E[x|y_2, y_e] - E[x|y_2]$, which by Lemma 2.4 is orthogonal to y_1 and y_2. The leader's strategy after the follower acquires y_e is the following (we omit the arguments in the strategies $\gamma_i^*(\cdot)$ and

$\gamma_1^0(\cdot)$ for a while to avoid the tedious expressions):

$$\gamma_1^0 = \arg\min_{\gamma_1(y_1)\in\Gamma_1} E\left\{\frac{1}{2}\gamma_1'\gamma_1 + \frac{1}{2}\gamma_2^{0'}P_1\gamma_2^0 + \gamma_1'Q_1\gamma_2^0 + \gamma_1'S_{11}x + \gamma_2^{0'}S_{12}x\right\}$$

$$= \arg\min_{\gamma_1(y_1)\in\Gamma_1} E\left\{\frac{1}{2}\gamma_1'\gamma_1 + \frac{1}{2}\gamma_2^{*'}P_1\gamma_2^* - \gamma_2^{*'}P_1S_{22}\hat{y} + \frac{1}{2}(S_{22}\hat{y})'P_1(S_{22}\hat{y}) + \gamma_1'Q_1\gamma_2^* \right.$$
$$\left. - \gamma_1'Q_1S_{22}\hat{y} + \gamma_1'S_{11}x + \gamma_2^{*'}S_{12}x - (S_{22}\hat{y})'S_{12}x\right\}$$

$$= \arg\min_{\gamma_1(y_1)\in\Gamma_1} E\left\{\frac{1}{2}\gamma_1'\gamma_1 + \frac{1}{2}\gamma_2^{*'}P_1\gamma_2^* + \gamma_1'Q_1\gamma_2^* + \gamma_1'S_{11}x + \gamma_2^{*'}S_{12}x\right\}$$

$$= \gamma_1^* \tag{59}$$

where we use the orthogonality conditions to get rid of the terms $\gamma_2^{*'}P_1S_{22}\hat{y}$ and $\gamma_1'Q_1S_{22}\hat{y}$ in taking the expectation operations. □

Theorem 5.3 If the follower acquires extra measurement y_e such that either one of Conditions $\overline{C}$ holds, then the follower can do better by incurring lower cost.

Proof: The proof is similar to Theorem 2.6 and hence omitted.

Now consider a feedback Stackelberg game with the same formulation as in the feedback Nash game of Section III except we consider two cases which correspond to two different information structures. Let $I_i(k)$ denote the information available to decision maker i at stage k, then

Case A: $I_1^A(k) = \{\hat{x}_1(k), \hat{x}_3(k)\}$, $I_2^A(k) = \{\hat{x}_1(k)\}$.

Case B: $I_1^B(k) = \{\hat{x}_1(k)\}$, $I_2^B(k) = \{\hat{x}_1(k), \hat{x}_3(k)\}$.

Let us call decision maker 1 the leader and 2 the follower. A pair (g_1^*, g_2^*) is a feedback Stackelberg solution to the game if

$$\sup_{\gamma_2^k \in R_k(^*\gamma_1^k)} J_1(g_1^*, g_{2k}^*, \gamma_2^k) \leq \sup_{\gamma_2^k \in R_k(\gamma_1^k)} J_1(g_{1k}^*, \gamma_1^k, g_{2k}^*, \gamma_2)$$

$\forall$ admissible γ_1^k. Where

$$g_{ik} \triangleq \{\gamma_i^0, \gamma_i^1, \ldots, \gamma_i^{k-1}, \gamma_i^{k+1}, \ldots, \gamma_i^{N-1}\}$$

$R_k(\gamma_1^k)$ is called the rational reaction set of the follower at stage k to the strategy γ_1^k announced by the leader and is defined by

$$R_k(\gamma_1^k) = \left\{ \hat{\gamma}_2^k \mid J_2(g_{1k}^*, \gamma_1^k, g_{2k}^*, \hat{\gamma}_2^k) \leq J_2(g_{1k}^*, \gamma_1^k, g_{2k}^*, \gamma_2^k) \quad \forall \text{ admissible } \gamma_2^k \right\}.$$

The feedback Stackelberg solution for Cases A and B are provided in Appendix C

Let Information I and II be defined as in Section III and satisfy the condition in Definition 4.1, then in Case A Information I provides better information for the leader alone than Information II while in Case B Information I provides better information for the follower alone than Information II. We have the following theorem.

Theorem 5.4 Under the information structure of Cases A and B, the feedback Stackelberg solution has the following properties:

(i) Better information for the leader alone is beneficial to the leader.

(ii) Better information for the follower alone is beneficial to the follower.

Proof: One way of proving this theorem is by using the connection of the feedback Stackelberg solution to the feedback Nash solution according to the procedure of [6] where it was proved that a feedback Stackelberg solution of an N-stage dynamic game is also a feedback Nash solution of a 2N-stage dynamic game and the result is then implied by Corollary 4.4 An independent proof of this theorem is provided in Appendix D.

Remark 5. A similar feedback Stackelberg game was studied in [8] where the expressions of the solution obtained were so complicated that it was not possible to investigate its informational properties. The expressions of the solution could have been simplified if the authors of [8] had observed the orthogonality condition in the information structure, i.e., Lemma 3.1 (ii) of this paper

Examples

Example 1 This example illustrates Theorem 2.5 and 2.6 under Condition C (i). Consider a static Nash game where all the notations follow those defined in Section II

$$J_1(\gamma_1, \gamma_2) = E[(x + u_1 + u_2)^2 + u_1^2]$$

$$J_2(\gamma_1, \gamma_2) = E[(x + u_1 + u_2)^2 + u_2^2]$$

Decision maker i has measurement y_i, $y_i = x + \omega_i$. x, ω_1 and ω_2 are independent random variables with zero mean and unit variance.

This example was previously considered in [2] and the Nash solution was given by $\gamma_1^*(y_1) = -\frac{1}{5} y_1$ and $\gamma_2^*(y_2) = -\frac{1}{5} y_2$ with corresponding Nash costs $J_1(\gamma_1^*, \gamma_2^*) = J_2(\gamma_1^*, \gamma_2^*) = \frac{468}{900}$. Now if in addition to y_2, decision maker 2 acquires extra measurement y_e, what is the impact to his Nash cost? It was shown (Case B of [2]) that if $y_e = y_1$ then decision maker 2 incurs higher Nash cost. In the following we will find a y_e such that $y_e \perp \{y_1, y_2\}$ and demonstrate that this y_e will lower decision maker 2's Nash cost.

Let $y_e = x - \omega_1 - \omega_2$, then it is easy to check that $y_e \perp \{y_1, y_2\}$. Denote the Nash solution after decision maker 2 acquires this y_e by (γ_1^0, γ_2^0), then by direct calculation we obtain

$$\gamma_1^0(y_1) = -\frac{1}{5} y_1$$

and

$$\gamma_2^0(y_2, y_e) = -\frac{1}{5} y_2 - \frac{1}{6} y_e$$

The corresponding Nash solution of decision maker 2 is $J_2(\gamma_1^0, \gamma_2^0)$ and

$$J_2(\gamma_1^0, \gamma_2^0) = \frac{318}{900} < \frac{468}{900} = J_2(\gamma_1^*, \gamma_2^*) \ .$$

<u>Example 2.</u> This example illustrates Corollary 4.3.
Consider a dynamic Nash game with the general formulation given in Section II and III. We choose $A = 0.5$, $\overline{x}_0 = 0$, $\Omega_0 = 10$, $B_i = P_i = R = 1$, $Q_i = 20$, $i = 1, 2$. Two kinds of information, I and II are described below:

Information I, $\hat{x}_1^I(\cdot)$, $\hat{x}_3^I(\cdot)$ are corresponding to

$$\left.\begin{aligned} y_1^I(\cdot) &= x(\cdot) + v_1^I(\cdot) \\ y_2^I(\cdot) &= x(\cdot) + v_2^I(\cdot) \end{aligned}\right. \qquad v_i^I(\cdot) \sim N(0,1),\ i=1,2 .$$

Information II, $\hat{x}_1^{II}(\cdot)$, $\hat{x}_3^{II}(\cdot)$ are corresponding to

$$\left.\begin{aligned} y_1^{II}(\cdot) &= x(\cdot) + v_1^{II}(\cdot) \\ y_2^{II}(\cdot) &= 0 \cdot x(\cdot) + v_2^{II}(\cdot) \end{aligned}\right. \qquad v_i^{II}(\cdot) \sim N(0,1),\ i=1,2 .$$

It is easy to see that for Information II, $\hat{x}_1^{II}(k) = \hat{x}_3^{II}(k)$ at every stage k and Information I provides better information for decision maker 1 alone than Information II. We compute the Nash cost of decision maker 1 for different number of stages, i.e., N from 1 to 19. The resulting costs are shown in Table 6.1. Notice that Information I is more beneficial to decision maker 1 than Information II. Two features of this fact are: first, it is independent of N, the number of stages and second, since $A = 0.5$, $\hat{x}_3^I(\cdot)$ is not dynamically better than $\hat{x}_3^{II}(\cdot)$.

	Information I	Information I	Benefit of Decision Maker 1 Due to Better Information for Him Alone
N= 1	16.72872	16.98826	0.259544
N= 2	19.79963	20.12271	0.323073
N= 3	21.68059	22.06824	0.387644
N= 4	23.31423	23.76004	0.445805
N= 5	24.90147	25.40363	0.502162
N= 6	26.48017	27.03831	0.558140
N= 7	28.05730	28.67135	0.614047
N= 8	29.63415	30.30409	0.669940
N= 9	31.21094	31.93677	0.725830
N=10	32.78773	33.56945	0.781720
N=11	34.36451	35.20212	0.837610
N=12	35.94123	36.83479	0.893500
N=13	37.51808	38.46747	0.949390
N=14	39.09486	40.10014	1.005280
N=15	40.67164	41.73281	1.061170
N=16	42.24843	43.36549	1.117060
N=17	43.82521	44.99816	1.172950
N=18	45.40199	46.63083	1.228840
N=19	46.97877	48.26350	1.284730

Table 6.1. Costs of decision maker 1 in Example 2 under different information versus different number of stages.

VI Conclusion

In a general two-decision-maker LQG Nash game (static or dynamic) we proved that more or better information for one of the decision makers alone is beneficial to him if he is informationally stronger than his opponent, i.e., he knows all his opponent's information. In a static game, more information to one of the decision makers alone is beneficial to him if such information is orthogonal to both decision makers' information. Such results are

quite understandable. Since Nash solution is an equilibrium solution with consistency constraint [9], any unilateral improvement of information does not guarantee benefit to either party. A unilateral improvement of information, however, does guarantee benefit to the

who has the improvement, if his opponent's strategy does not change by such improvement such that he who has the improved information can use it to optimize his strategy without constraint. In order that his opponent's strategy does not change, his opponent should be totally ignorant of this improved information and which is implied by the orthogonality condition given by Lemma 2.4. Similar results hold in static and feedback Stackelberg games for both the leader and the follower. The leader in a static Stackelberg game, however, can use any extra information to his benefit.

As we noted before, the investigation of the informational property of the dynamic Nash game is greatly simplified by the formulation of the game where a sort of separation principle holds and the estimation error is independent of the controls. Without these nice properties, it will be difficult either in defining "better information for one decision maker alone" or in solving for the Nash solution. Either one of the difficulties makes the problem extremely hard.

An extension of the results obtained in this paper to N-decision-maker Nash game is straight forward and such results constitute a fundamental step in designing information structure [10, 11, 12] for large scale systems.

APPENDIX A

Proof of Lemma 3.1.

Consider the following state equation and measurements:

$$\bar{x}(k+1) = A\bar{x}(k) + \omega(k) , \qquad \bar{x}(0) = x_0 \tag{A1}$$

$$\bar{y}(k) = H_i\bar{x}(k) + v_i(k) , \qquad i = 1, 2 . \tag{A2}$$

where x_0, $\{\omega(k)\}$ and $\{v_i(k)\}$ are defined as in Section III. By comparing (A1) with (17) we immediately have

$$x(k) = \bar{x}(k) + \sum_{n=0}^{k-1} A^{k-n-1}[B_1u_1(n) + B_2u_2(n)] \tag{A3}$$

Let

$$\hat{\bar{x}}_1(k) \triangleq E[\bar{x}(k) \mid \bar{y}_1(0), \ldots, \bar{y}_1(k)] \tag{A4}$$

and

$$\hat{\bar{x}}_3(k) \triangleq E[\bar{x}(k) \mid \bar{y}_1(0), \ldots, \bar{y}_1(k), \bar{y}_2(0), \ldots, \bar{y}_2(k)] \tag{A5}$$

then $\hat{\bar{x}}_i(k)$, $i = 1, 3$ are given exactly by the Kalman filter equations (19) except that (19b) is replaced by

$$\hat{\tilde{x}}_i(k+1/k) = A\hat{\tilde{x}}_i(k) \tag{A6}$$

By the construction of $\hat{\tilde{x}}_i(k)$ and $\hat{x}_i(k)$, $i=1,3$, it is easy to see that

$$\hat{x}_i(k) = \hat{\bar{x}}_i(k) + \varphi_k \qquad (A7)$$

where

$$\varphi_k = \sum_{n=0}^{k-1} A^{k-n-1}[B_1u_1(n) + B_2u_2(n)] \qquad (A8)$$

Since $\hat{\bar{x}}_3(k)$ is a refinement of $\hat{\bar{x}}_1(k)$, we obtain

$$\begin{aligned}\hat{\bar{x}}_1(k) &= E[\bar{x}(k) \mid \bar{y}_1(0),\ldots,\bar{y}_1(k)] \\ &= E[E[\bar{x}(k) \mid \bar{y}_1(0),\ldots,\bar{y}_1(k),\bar{y}_2(0),\ldots,\bar{y}_2(k)] \mid \bar{y}_1(0),\ldots,\bar{y}_1(k)] \\ &= E[\hat{\bar{x}}_3(k) \mid \bar{y}_1(0),\ldots,\bar{y}_1(k)] \qquad (A9)\end{aligned}$$

Hence

$$\begin{aligned}E[\hat{\bar{x}}_3(k) \mid \hat{\bar{x}}_1(k)] &= E[E[\hat{\bar{x}}_3(k) \mid \bar{y}_1(0),\ldots,\bar{y}_1(k)] \mid \hat{\bar{x}}_1(k)] \\ &= E[\hat{\bar{x}}_1(k) \mid \hat{\bar{x}}_1(k)] = \hat{\bar{x}}_1(k) \;. \qquad (A10)\end{aligned}$$

(A7) indicates that

$$\begin{aligned}E[\hat{x}_3(k) \mid \hat{x}_1(k)] &= E[\hat{\bar{x}}_3(k) + \varphi_k \mid \hat{\bar{x}}_1(k) + \varphi_k] \\ &= E[\hat{\bar{x}}_3(k) \mid \hat{\bar{x}}_1(k)] + \varphi_k = \hat{\bar{x}}_1(k) + \varphi_k \\ &= \hat{x}_1(k) \qquad (A11)\end{aligned}$$

By the projection theorem [4], $\hat{x}_3(k) - E[\hat{x}_3(k) \mid \hat{x}_1(k)]$ is of zero mean and orthogonal to $\hat{x}_1(k)$, i.e., $E[\hat{x}_4(k)] = 0$ and $\hat{x}_4(k) \perp \hat{x}_1(k)$. Finally

$$E[\hat{x}_3(k)\ \hat{x}_3'(k)] = E[(\hat{x}_1(k) + \hat{x}_4(k))(\hat{x}_1(k) + \hat{x}_4(k))']$$

$$= E[\hat{x}_1(k)\ \hat{x}_1'(k)] + E[\hat{x}_4(k)\ \hat{x}_4'(k)] \qquad (A12)$$

i.e.,

$$E[\hat{x}_4(k)\ \hat{x}_4'(k)] = E[\hat{x}_3(k)\ \hat{x}_3'(k)] - E[\hat{x}_1(k)\ \hat{x}_1'(k)]$$

$$= \Sigma_1(k) - \Sigma_3(k) \qquad (A13)$$

APPENDIX B

In this appendix we prove Theorem 3.2. The proof is similar to that of Theorem 2.1 of [1]. Since the Nash solution g_i^* of decision maker i is a solution of the optimal control problem where the decision maker j, $j \neq i$ fixes his strategy at g_j^*, we can solve the problem by dynamic programming. Recall that $J_i(k)$ denotes the cost to go of decision maker i at stage k.

At stage N,

$$J_i(N) = E[x'(N)P_i x(N)] = E[\hat{x}_3'(N)P_i\hat{x}_3(N)] + tr[P_i\Sigma_3(N)]$$

$$= E[\hat{x}_3'(N)L_i(N)\hat{x}_3(N)] + K_i(N) \qquad (B1)$$

$$i = 1, 2$$

where

$$L_i(N) \triangleq P_i, \quad K_i(N) \triangleq tr\,[P_i\Sigma_3(N)] .$$

At stage N-1

$$J_i(N-1) = E[x'(N-1)P_i x(N-1) + u_i'(N-1)u_i(N-1) + u_j'(N-1)Q_i u_j(N-1) +$$

$$x(N)P_i x(N)] \qquad j \neq i, \quad i,j = 1,2 . \qquad (B2)$$

After receiving $I_i(N-1)$, decision maker i's objective is to minimize $\overline{J}_i(N-1)$ given by

$$\overline{J}_i(n-1) = E[x'(N-1)P_i x(N-1) + u_i'(N-1)u_i(N-1) + u_j'(N-1)Q_i u_j(N-1)$$
$$+ x'(N)P_i x(N) \mid I_i(N-1)] \qquad \text{(B3)}$$

By applying the Kalman filter equations (19) and Lemma 4.3.1 we obtain

$$\overline{J}_1(N-1) = u_1'(N-1)u_1(N-1) + u_2'(N-1)Q_1 u_2(N-1) + (A\hat{x}_3(N-1) + B_1 u_1(N-1)$$
$$+ B_2 u_2(N-1))' L_1(N)(A\hat{x}_3(N-1) + B_1 u_1(N-1) + B_2 u_2(N-1))$$
$$+ \hat{x}_3'(N-1)P_1\hat{x}_3(N-1) + \mathrm{tr}\{P_1\Sigma_3(N-1) + L_1(N)[\Sigma_3(N/N-1) - \Sigma_3(N)]\}$$
$$+ K_1(N) \qquad \text{(B4)}$$

and

$$\overline{J}_2(N-1) = E[(A\hat{x}_3(N-1)+B_1u_1(N-1)+B_2u_2(N-1))' L_2(N)(A\hat{x}_3(N-1)$$
$$+ B_1u_1(N-1)+B_2u_2(N-1)) \mid I_2(N-1)]$$
$$+ E[u_1'(N-1)Q_2u_1(N-1) \mid I_2(N-1)] + \hat{x}_1'(N-1)P_2\hat{x}_1(N-1)$$
$$+ \mathrm{tr}\{P_1\Sigma_1(N-1) + L_2(N)[\Sigma_3(N/N-1) - \Sigma_3(N)]\} + K_2(N) \qquad \text{(B5)}$$

Since $\overline{J}_i(N-1)$ is convex in $u_i(N-1)$, the Nash pair at stage N-1, ${}^*\gamma_1^{N-1}$, ${}^*\gamma_2^{N-1}$ is chosen such that

$$\left.\frac{\partial\overline{J}_i(N-1)}{\partial\gamma_i^{N-1}}\right|_{{}^*\gamma_1^{N-1},\ {}^*\gamma_2^{N-1}} = 0, \quad i = 1,2 \qquad \text{(B6)}$$

We then have

$$ {}^{*}\gamma_1^{N-1}(I_1(N-1)) = -[I+B_1'L_1(N)B_1]^{-1}B_1'L_1(N)[A\hat{x}_3(N-1)+B_2\,{}^{*}\gamma_2^{N-1}(I_2(N-1))] \tag{β7} $$

$$ {}^{*}\gamma_2^{N-1}(I_2(N-2)) = -[I+B_2'L_2(N)B_2]^{-1}B_2'L_2(N)[A\hat{x}_1(N-1)+B_1E[{}^{*}\gamma_1^{N-1}(I_1(N-1)) \;|\; I_2(N-1)]] \tag{β8} $$

From (β7) and by Lemma 3.1 (i) we obtain

$$ E[{}^{*}\gamma_1^{N-1}(I_1(N+1) \,|\, I_2(N-1)] = -[I+B_1'L_1(N)B_1]^{-1}B_1'L_1(N)[A\hat{x}_1(N-1) + B_2\,{}^{*}\gamma_2^{N-1}(I_2(N-1))] \tag{β9} $$

Substituting (β9) into (β8),

$$ {}^{*}\gamma_2^{N-2}(I_2(N-1)) = -[I+B_2'L_2(N)B_2]^{-1}B_2'L_2(N)[A\hat{x}_1(N-1) - B_1(I+B_1'L_1(N)B_1)^{-1} B_1'L_1(N)(A\hat{x}_1(N-1) + B_2\,{}^{*}\gamma_2^{N-1}(I_2(N-1))] \tag{β10} $$

By applying the following formula (β11) several times we obtain (β12).

$$ Z_1(I + Z_2Z_1) = (I + Z_1Z_2)^{-1}Z_1 \tag{β11} $$

$$ \begin{aligned} {}^{*}\gamma_2^{N-1}(I_2(N-1)) &= -B_2'L_2(N)[I+B_1B_1'L_1(N) + B_2B_2'L_2(N)]^{-1}A\hat{x}_1(N-1) \\ &= F_2(N-1)\hat{x}_1(N-1) \end{aligned} \tag{β12} $$

where

$$ F_2(N-1) \triangleq -B_2'L_2(N)[I+B_1B_1'L_1(N) + B_2B_2'L_2(N)]^{-1}A \tag{β13} $$

Substituting (β12) into (β7) we obtain

$$^*\gamma_1^{N-1}(I_1(N-1)) = -B_1'L_1(N)[I+B_1B_1'L_1(N)]^{-1}A\hat{x}_4(N-1) - B_1'L_1(N)[I+B_1B_1'L_1(N) + B_2B_2'L_2(N)]^{-1}A\hat{x}_1(N-1)$$

$$= F_{11}(N-1)\hat{x}_1(N-1) + F_{14}(N-1)\hat{x}_4(N-1) \qquad (\beta 14)$$

where

$$F_{11}(N-1) \triangleq -B_1'L_1(N)[I+B_1B_1'L_1(N)+B_2B_2'L_2(N)]^{-1}A \qquad (\beta 15)$$

and

$$F_{14}(N-1) \triangleq -B_1'L_1(N)[I+B_1B_1'L_1(N)]^{-1}A \qquad (\beta 16)$$

Notice that $(^*\gamma_1^{N-1}, {}^*\gamma_2^{N-1})$ given by (β12) and (β14) exists and is unique if $[I + B_1B_1'L_1(N) + B_2B_2'L_2(N)]$ is nonsingular.

Substituting (β12) and (β14) into (β2) we obtain

$$J_i(N-1) = E[\hat{x}_3'(N-1)L_i(N-1)\hat{x}_3(N-1)] + K_i(N-1) \qquad (\beta 17)$$

where $L_i(N-1)$ and $K_i(N-1)$ are given by (26) and (33) respectively. As we can see, (β17) and (β1) are of the same form. In deriving the Nash pair $(^*\gamma_1^{N-2}, {}^*\gamma_2^{N-2})$ at stage N-2, we will repeat what we did at stage N-1. An inductive argument then proves the theorem.

APPENDIX C

In this appendix we derive the feedback Stackelberg solution, the problem was stated in Section VI.

<u>Theorem C:</u> There exists a unique solution to the feedback Stackelberg game, (i) the solution for Case A is

$$u^*_{1A}(k) = {}^*\gamma^k_{1A}(I^A_1(k)) = F_{11A}(k)\hat{x}_1(k) + F_{14A}(k)\hat{x}_4(k) \tag{C1a}$$

$$\begin{aligned} u^*_{2A}(k) = {}^*\gamma^k_{2A}(I^A_2(k)) &= F_{2A}(k)\{A\hat{x}_1(k)+B_1E[{}^*\gamma^k_{1A}(I^A_1(k)) \mid I^A_2(k)]\} \\ &= F_{21A}(k)\hat{x}_1(k) \end{aligned} \tag{C1b}$$

where

$$F_{11A}(k) = -B_1'Z_A(k+1)[I+B_1B_1'Z_A(k+1)]^{-1}A \tag{C2}$$

$$F_{14A}(k) = -B_1'L_{1A}(k+1)[I+B_1B_1'L_{1A}(k+1)]^{-1}A \tag{C3}$$

$$F_{2A}(k) = -B_2'L_{2A}(k+1)[I+B_2B_2'L_{2A}(k+1)]^{-1} \tag{C4}$$

$$F_{21A}(k) = -B_2'L_{2A}(k+1)[I+B_2B_2'L_{2A}(k+1)]^{-1}[I+B_1B_1'Z_A(k+1)]^{-1}A \tag{C5}$$

$$\begin{aligned} Z_A(k) = [I+B_2B_2'L_{2A}(k)]^{-1'}[L_{2A}(k)B_2Q_1B_2'L_{2A}(k)+L_{1A}(k)] \\ [I+B_2B_2'L_{2A}(k)]^{-1} \end{aligned} \tag{C6}$$

$$L_{1A}(k) = P_1+F'_{11A}(k)F_{11A}(k) + F'_{21A}(k)Q_1F_{21A}(k) +$$

$$(A+B_1F_{11A}(k)+B_2F_{21A}(k))'L_{1A}(k+1)(A+B_1F_{11A}(k)+B_2F_{21A}(k)),$$

$$L_{1A}(N) = P_1 \quad . \tag{C7}$$

$$L_{2A}(k) = P_2+F'_{11A}(k)Q_2F_{11A}(k)+F'_{21A}(k)F_{21A}(k) +$$

$$(A+B_1F_{11A}(k)+B_2F_{21A}(k))'L_{2A}(k+1)(A+B_1F_{11A}(k)+B_2F_{21A}(k)),$$

$$L_{2A}(N) = P_2 \, . \tag{C8}$$

Their costs to go at stage k are respectively

$$J_{1A}(k) = E\{\hat{x}_3'(k)L_{1A}(k)\hat{x}_3(k)\} + K_{1A}(k) \tag{C9}$$

$$J_{2A}(k) = E\{\hat{x}_4'(k)L_{2A}(k)\hat{x}_3(k)\} + K_{2A}(k) \tag{C10}$$

where

$$K_{1A}(k) = tr\{[P_1+A'L_{1A}(k+1)A-L_{14A}(k)]\Sigma_3(k) - L_{1A}(k+1)\Sigma_3(k+1) +$$
$$L_{14A}(k)\Sigma_1(k)+L_{1A}(k+1)R\}+K_{1A}(k+1), \; K_{1A}(N) = tr\{P_1\Sigma_3(N)\}. \tag{C11}$$

$$K_{2A}(k) = tr\{[P_2+A'L_{2A}(k+1)A-L_{24A}(k)]\Sigma_3(k)- L_{2A}(k+1)\Sigma_3(k+1) +$$
$$L_{24A}(k)\Sigma_1(k)+L_{2A}(k+1)R\}+K_{2A}(k+1), \; K_{2A}(N) = tr\{P_2\Sigma_3(N)\}. \tag{C12}$$

$$L_{14A}(k) = P_1+F'_{14A}(k)F_{14A}(k)+(A+B_1F_{14A}(k))'L_{1A}(k+1)(A+B_1F_{14A}(k))-L_{1A}(k) \tag{C13}$$

$$L_{24A}(k) = P_2+F'_{14A}(k)Q_2F_{14A}(k)+(A+B_1F_{14A}(k))'L_{2A}(k+1)(A+B_1F_{14A}(k)-L_{2A}(k) \tag{C14}$$

(ii) The solution for Case B is

$$u^*_{1B}(k) = {}^*\gamma^k_{1B}(I^B_1(k)) = F_{11B}(k)\hat{x}_1(k) \tag{C15a}$$

$$u^*_{2B}(k) = {}^*\gamma^k_{2B}(I^B_2(k)) = F_{2B}(k)\{A\hat{x}_3(k)+B_1{}^*\gamma^k_{1B}(I^B_1(k))\}$$

$$= F_{21B}(k)\hat{x}_1(k) + F_{24B}(k)\hat{x}_4(k) \tag{C15b}$$

where

$$F_{11B}(k) = -B'_1Z_B(k+1)[I + B_1B'_1Z_B(k+1)]^{-1}A \tag{C16}$$

$$F_{2B}(k) = -B'_2L_{2B}(k+1)[I + B_2B'_2L_{2B}(k+1)]^{-1} \tag{C17}$$

$$F_{21B}(k) = -B'_2L_{2B}(k+1)[I+B_2B'_2L_{2B}(k+1)]^{-1}[I+B_1B'_1Z_B(k+1)]^{-1}A \tag{C18}$$

$$F_{24B}(k) = -B'_2L_{2B}(k+1)[I+B_2B'_2L_{2B}(k+1)]^{-1}A \tag{C19}$$

$$Z_B(k) = [I+B_2B'_2L_{2B}(k)]^{-1'}[L_{2B}(k)B_2Q_1B'_2L_{2B}(k)+L_{1B}(k)]$$
$$[I+B_2B'_2L_{2B}(k)]^{-1} \tag{C20}$$

$$L_{1B}(k) = P_1+F'_{11B}(k)F_{11B}(k)+F'_{21B}(k)Q_1F_{21B}(k) +$$
$$(A+B_1F_{11B}(k)+B_2F_{21B}(k))'L_{1B}(k+1)(A+B_1F_{11B}(k)+B_2F_{21B}(k)),\ L_{1B}(N)=P_1 . \tag{C21}$$

$$L_{2B}(k) = P_2+F'_{11B}(k)Q_2F_{11B}(k)+F'_{21B}(k)F_{21B}(k) +$$

$$(A+B_1F_{11B}(k)+B_2F_{21B}(k))'L_{2B}(k+1)(A+B_1F_{11B}(k)+B_2F_{21B}(k)) ,$$

$$L_{2B}(N) = P_2 \quad . \tag{C22}$$

Their costs to go at stage k are respectively

$$J_{1B}(k) = E\{ \hat{x}'_3(k)L_{1B}(k)\hat{x}_3(k)\} + K_{1B}(k) \tag{C23}$$

$$J_{2B}(k) = E\{ \hat{x}'_3(k)L_{2B}(k)\hat{x}_3(k)\} + K_{2B}(k) \tag{C24}$$

where

$$K_{1B}(k) = tr\{[P_1+A'L_{1B}(k+1)A-L_{14B}(k)]\Sigma_3(k)-L_{1B}(k+1)\Sigma_3(k+1) +$$

$$L_{14B}(k)\Sigma_1(k)+L_{1B}(k+1)R\} + K_{1B}(k+1), \; K_{1B}(N) = tr\{P_1\Sigma_3(N)\} . \tag{C25}$$

$$K_{2B}(k) = tr\{[P_2+A'L_{2B}(k+1)A-L_{24B}(k)]\Sigma_3(k)-L_{2B}(k+1)\Sigma_3(k+1) +$$

$$L_{24B}(k)\Sigma_1(k)+L_{2B}(k+1)R\}+K_{2B}(k+1), \; K_{2B}(N) = tr\{P_2\Sigma_3(N)\}. \tag{C26}$$

$$L_{14B}(k) = P_1+F'_{24B}(k)Q_1F_{24}(k)+(A+B_2F_{24B}(k))'L_{1B}(k+1)(A+B_2F_{24B}(k)-L_{1B}(k). \tag{C27}$$

$$L_{24B}(k)=P_2+F'_{24B}(k)F_{24B}(k)+(A+B_2F_{24B})'L_{2B}(k+1)(A+B_2F_{24B}(k)-L_{2B}(k). \tag{C28}$$

Remark: *It is easy to see that in the above theorem*

$$F_{11B}(k) = F_{11A}(k)\ , \qquad Z_B(k) = Z_A(k)$$
$$F_{2B}(k) = F_{2A}(k)\ , \qquad L_{1B}(k) = L_{1A}(k)$$
$$F_{21B}(k) = F_{21A}(k)\ , \qquad L_{2B}(k) = L_{2A}(k)$$

Proof of Theorem C: We will prove part (i) only, the proof for part (ii) is similar.

Feedback Stackelberg strategies have the property that they are in static Stackelberg equilibrium at every stage of the problem. This property can be observed from its definition and hence we can solve the problem by going backwards (a dynamic programming type of approach).

At stage N (no more decisions to be made), the cost to go of decision maker i is

$$\begin{aligned} J_i(N) &= E[x'(N)P_i x(N)] \\ &= E[\hat{x}_3'(N)P_i\hat{x}_3(N)] + tr\{P_i\Sigma_3(N)\} \\ &= E[\hat{x}_3'(N)L_{iA}(N)\hat{x}_3(N)] + K_{iA}(N) \end{aligned} \tag{C.29}$$

where

$$L_{iA}(N) = P_i\ , \qquad K_{iA}(N) = tr\{P_i\Sigma_3(N)\}$$

At stage N-1 ($I_i^A(N-1)$ is available), decision maker i's objective is to minimize $\overline{J}_i(N-1)$ given by

$$\overline{J}_i(N-1) = E[x'(N-1)P_i x(N-1)+u_i'(N-1)u_i(N-1)+u_j'(N-1)Q_i u_j(N-1)+x'(N)P_i x(N) \\ |I_i(N-1)] \tag{C.30}$$

By applying the Kalman filter equations (19) and Lemma 3.1 we obtain

$$\bar{J}_1(N-1) = u_1'(N-1)u_1(N-1)+u_2'(N-1)Q_1u_2(N-1)+(A\hat{x}_3(N-1)+B_1u_1(N-1)$$
$$+B_2u_2(N-1))'L_{1A}(N)(A\hat{x}_3(N-1)+B_1u_1(N-1)+B_2u_2(N-1))$$
$$+\hat{x}_3'(N-1)P_1\hat{x}_3(N-1)+\mathrm{tr}\{P_1\Sigma_3(N-1)+L_{1A}(N)[\Sigma_3(N/N-1)-\Sigma_3(N)]\}$$
$$+K_{1A}(N) \tag{C31}$$

and

$$\bar{J}_2(N-1) = E[(A\hat{x}_3(N-1)+B_1u_1(N-1)+B_2u_2(N-1))'L_{2A}(N)(A\hat{x}_3(N-1)+B_1u_1(N-1)$$
$$+B_2u_2(N-1))\,|I_2(N-1)]+E[u_1'(N-1)Q_2u_1(N-1)\,|I_2(N-1)]$$
$$+\hat{x}_1(N-1)P_2\hat{x}_1(N-1)+\mathrm{tr}\{P_1\Sigma_1(N-1)+L_{2A}(N-1)[\Sigma_3(N/N-1)-\Sigma_3(N)]\}$$
$$+K_{2A}(N) \tag{C32}$$

To any strategy $\gamma_{1A}^{N-1}(I_1^A(N-1))$ announced by the leader, the follower's rational reaction set is a singleton, i.e.,

$$\gamma_{2A}^{N-1}(I_2^A(N-1)) = -B_2'L_{2A}[I+B_2B_2'L_{2A}(N)]^{-1}[A\hat{x}_1(N-1)$$
$$+B_1E[\gamma_{1A}^{N-1}(I_1^A(N-1)\,|I_2(N-1)] \tag{C33}$$

Substituting $u_2(N-1)$ given by (C33) into (C31) and optimizing $\bar{J}_1(N-1)$ with respect to $u_1(N-1)$ we obtain

$$u_1^*(N-1) = F_{11A}(N-1)\hat{x}_1(N-1)+F_{14A}(N-1)\hat{x}_4(N-1) \tag{C34}$$

where $F_{11A}(N-1)$ and $F_{14A}(N-1)$ are given respectively by (C2) and (C3). Substituting (C34) into (C33) we obtain $u_2^*(N-1)$ given by (C16).

Substituting $u_1^*(N-1)$ and $u_2^*(N-1)$ into $J_i(N-1)$ we obtain (H9) and (C10) for $k = N-1$. The proof of this feedback Stackelberg solution can then be concluded by an inductive argument.

APPENDIX D

In this appendix we prove Theorem 5.4. We will prove part (i) only, the proof for part (ii) is similar.

From equation (C9) and (C11) of Appendix C we obtain that the cost for the leader in Case A is

$$J_{1A}(0) = \mathrm{tr}\{L_{1A}(0)\hat{\Omega}_3^I(0) + [P_i + A'L_{1A}(1)A - L_{14A}(0)]\Sigma_3(0) + L_{1A}(1)R + L_{14A}(0)\Sigma_1(0) + \sum_{k=1}^{N-1}[[P_1 + A'L_{1A}(k+1)A - L_{1A}(k) - L_{14A}(k)]\Sigma_3(k) + L_{1A}(k+1)R + L_{14A}(k)\Sigma_1(k)]\} \tag{D1}$$

Let $J_{1A}^I(0)$ and $J_{1A}^{II}(0)$ correspond to Information I and II respectively, then

$$J_{1A}^{II}(0) - J_{1A}^{I}(0) = \sum_{k=0}^{N-1} \mathrm{tr}\{[P_1 + A'L_{1A}(k+1)A - L_{1A}(k) - L_{14A}(k)][\Sigma_3^{II}(k) - \Sigma_3^{I}(k)] + L_{14A}(k)[\Sigma_1^{II}(k) - \Sigma_1^{I}(k)]\} \tag{D2}$$

If Information I provides better information for the leader alone than Information II, then Lemma 4.2 implies $J_{1A}^{II}(0) \geqq J_{1A}^{I}(0)$ if

$$P_1 + A'L_{1A}(k+1)A - L_{1A}(k) - L_{14A}(k) \geqq 0 \quad \text{for every } k \in \theta \ .$$

Substituting equation (c 7) and (C 13) of Appendix C into the left hand side of the above equation we obtain

$$P_1 + A' L_{1A}(k+1)A - L_{1A}(k) - L_{14}(k) = \{ B_1' L_{1A}(k+1)[I + B_1 B_1' L_{1A}(k+1)]^{-1} A \}'$$

$$[I + B_1' L_{1A}(k+1) B_1] B_1' L_{1A}(k+1)[I + B_1 B_1' L_{1A}(k+1)]^{-1} A$$

$$\geq 0 . \qquad \text{(D3)}$$

REFERENCE

[1] M. Tu and G. P. Papavassilopoulos "On the Informational properties of the Nash Solution of LQG Dynamic Games" to appear on IEEE Trans Aut. Control.

[2] Y. C. Ho, I. Blau and T. Basar, "A Tale of Four Information Structures," Lecture Notes in Economics and Mathematical Systems," Vol. 107, Springer-Verlag, 1975.

[3] T. Basar, "Equilibrium Solutions in Two-Person Quadratic Decision Problems with Static Information Structures," IEEE Transactions on Automatic Control, Vol. AC-20, No. 3, June 1975.

[4] B.D.O. Anderson and J. B. Moore, "Optimal Filtering," Prentice-Hall, 1979.

[5] G. P. Papavassilopoulos, "Solution of Some Stochastic Quadratic Nash and Leader-Follower Games," SIAM J. Control and Optimization, Vol. 19, No. 5, September 1981.

[6] T. Basar and A. Haurie, "Feedback Equilibria in Differential Games with Structural and Modal Uncertainties," J. B. Cruz, Ed., Advances in Large Scale Systems, JAC Publica., 1983.

[7] T. Basar, "Stochastic Stagewise Stackelberg Strategies for Linear Quadratic Systems," Lecture Notes in Control and Information Sciences, Vol. 16, 1979.

[8] D. Castanon and M. Athans, "On Stochastic Dynamic Stackelberg Strategies," Automatica, Vol. 12, pp. 177-183.

[9] P. M. Walsh and J. B. Cruz, Jr., "An Example of the Impact of Information Structure on Decentralized Multicriterion Control Problems," IEEE Transactions on Automatic Control, Vol. AC-24, December, 1979.

[10] K. C. Chu, "Designing Information Structures for Quadratic Decision Problems," Journal of Optimization Theory and Applications, Vol. 25, No. 1, May 1978.

[11] P. M. Walsh, "Decentralized Control and Multicriterion Decision Making," Ph.D. Thesis, University of Illinois at Urbana-Champaign, Dept. of Electrical Engineering, 1980.

[12] Y. C. Ho and K. C. Chu, "Information Structure in Dynamic Multi-Person Control Problems," Automatica, Vol. 10, pp. 341-351, Pergamon Press, 1974.

2

Impact of explicit and implicit control sharing on the performance of two-person one-act LQG Nash games*

M. TU** and G.P. PAPAVASSILOPOULOS

Department of Electrical Engineering – Systems, University of Southern California, Los Angeles, CA 90089, U.S.A.

Received
Revised 21 March 1984
Communicated by A.P. Sage

A two-person one-act LQG Nash game is considered under three different information structures: explicit control sharing, implicit control sharing and static information. The relations among the corresponding solutions and their impacts on the resulting costs are studies.

1. Introduction

In game problems, the players have certain kinds of information; they make decisions based on this information. We say that there is explicit control sharing (ECS) in a game if a player's information includes the previous control values of other players. Two previous works concerning the impact of ECS on the optimal costs in Nash games were reported in [1] and [2]. In [1] a two-person LQG Nash game was considered where the information structure is partially nested and each player acts once and it was shown (theorem 2 of [1]) that the first player might do better if he reveals his control value to the second player than he could do in a static information structure (SIS). It is known that in Nash games, if there is ECS then in general there exist many solutions [8]. Uchida considered an example of a two-person LQG Nash game [2] where the information is partially nested and each player acts once, and showed that among the nonunique solutions under ECS, one of them is equivalent to the SIS solution. Furthermore, it is claimed in [2] that this SIS solution gives a local minimum of the first player's cost among the linear class of the nonunique solutions. In other words, the first player might do better at least locally in a SIS than if he reveals his control value to the second player. The claim which Uchida did not prove and the result of Ho, Blau and Basar in [1] seem to contradict each other.

In this paper we consider a two-person LQG Nash game where the information is partially nested and each player acts once. We study the impact that the first player, who reveals his control value explicitly and implicitly to the second player, has on the first player's Nash cost. By implicit control sharing (ICS) we mean that player 2 has a noise-corrupted measurement which is affine in the system state and player 1's control. Our aim is to relate the Nash solutions under ECS to those under ICS and give a full view of

* This research was supported by AFOSR grant 82-0174 and by the USC Faculty Research and Innovation Fund.
** Current address: Bell Communications Research, 331 Newman Spring Road, NVC-1F255, Red Bank, NJ 07716, U.S.A.

North-Holland
Large Scale Systems 7 (1984) 219–226

0167-420X/84/$3.00

solutions exist under certain nonsingularity conditions. Here we state the solutions and for proofs we refer to [6]–[8]. The impact of ICS and ECS is then considered by comparing the Nash costs $J_1(\gamma_1, \gamma_2)$ of cases A and B (including the Stackelberg cost) with that of Case C.

2.1. Nash solution of Case A

Under the condition

$$1 + q_1 b_1^2 + q_2 b_2^2 + \zeta_A (q_1 - p_{12} q_2) b_1 b_2 \neq 0 \tag{7}$$

the unique Nash solution of Case A is given by

$$\gamma_{1A}(y_1) = -\{1 + q_1 b_1^2 + q_2 b_2^2 + \zeta_A (q_1 - p_{12} q_2) b_1 b_2\}^{-1} \{b_1 q_1 + \zeta_A (q_1 - p_{12} q_2) b_2\} \cdot h_1 (1 + h_1^2)^{-1} y_1 , \tag{8a}$$

$$\gamma_{2A}(y_1, y_2) = -(1 + q_2 b_2^2)^{-1} q_2 b_2 \{b_1 \gamma_{1A}(y_1) + [h_1 y_1 + h_2 (y_2 - d\gamma_{1A}(y_1))] \cdot (1 + h_1^2 + h_2^2)^{-1}\} , \tag{8b}$$

where

$$\zeta_A = -(1 + q_2 b_2^2)^{-1} q_2 b_2 h_2 d (1 + h_1^2 + h_2^2)^{-1} . \tag{9}$$

Notice that $(\gamma_{1A}, \gamma_{2A})$ depends on ζ_A, which in turn depends on d. To different d's, corresponds different pairs $(\gamma_{1A}, \gamma_{2A})_d$ provided that (7) holds. Let us call M the class of all these solutions $(\gamma_{1A}, \gamma_{2A})_d$ for varying values of d.

2.2. Nash solution of Case B (linear class)

There exist uncountably many Nash solutions for Case B, with the linear ones given by:

$$\gamma_{1B}(y_1) = -\{1 + q_1 b_1^2 + q_2 b_2^2 + \zeta (q_1 - p_{12} q_2) b_1 b_2\}^{-1} \cdot \{b_1 q_1 + \zeta (q_1 - p_{12} q_2) b_2\} \cdot h_1 (1 + h_1^2)^{-1} y_1 , \tag{10a}$$

$$\gamma_{2B}(y_1, y_2^0, u_1) = -(1 + q_2 b_2^2)^{-1} q_2 b_2 \{b_1 \gamma_{1B}(y_1) + [h_1 y_1 + h_2 y_2^0](1 + h_1^2 + h_2^2)^{-1}\} + \zeta (u_1 - \gamma_{1B}(y_1)) , \tag{10b}$$

where ζ is any real number such that

$$1 + q_1 b_1^2 + q_2 b_2^2 + \zeta (q_1 - p_{12} q_2) b_1 b_2 \neq 0 . \tag{11}$$

Let us denote by L the class of all these linear solutions $(\gamma_{1B}, \gamma_{2B})_\zeta$.

2.3. Stackelberg solution of Case B

The Stackelberg solution with player 1 as the leader is denoted by $(\gamma_{1S}, \gamma_{2S})$ and is the following:

$$(\gamma_{1S}, \gamma_{2S}) = (\gamma_{1B}, \gamma_{2B})_{\zeta = \zeta_S} , \tag{12}$$

where

$$\zeta_S = -(1 + q_2 b_2^2)^{-1} q_2 b_1 b_2 . \tag{13}$$

2.4. Nash solution of Case C

The Nash solution of Case C is a special one of Case A with $\zeta_A = 0$ in (8). It is also a special one of Case B with $\zeta = 0$ in (10). Notice that (7) and (11) are satisfied when $\zeta_A = \zeta = 0$.

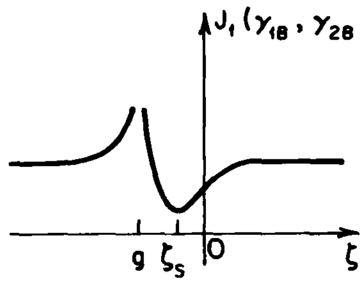


Fig. 1. Impact of ECS: $J_1(\gamma_{1B}, \gamma_{2B})$ as a function of ζ where ζ_S denotes the Stackelberg solution and 0 denotes the SIS Nash solution.

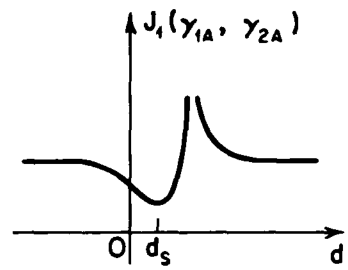


Fig. 2. Impact of ICS: $J_1(\gamma_{1A}, \gamma_{2A})$ as a function of d where $d_S = (1 + h_1^2 + h_2^2)b_1/h_2$ and 0 denotes the SIS Nash solution.

Theorem 1.

(i) *In L, the set of uncountably many linear Nash solutions under ECS, the unique local and global minimum of J_1 is given by $(\gamma_{1B}, \gamma_{2B})_{\zeta=\zeta_S}$ which is the Stackelberg solution.*

(ii) *Under ECS, player 1 can do better than under SIS if*

$$\zeta \in [\zeta_S, 0).$$

(iii) *Under ICS, player 1 can do better than under SIS if*

$$d \in (0, (1 + h_1^2 + h_2^2)b_1/h_2].$$

Remark 1. This theorem shows that Ucnida's claim, namely that the SIS solution is a local minimum of J_1 in L, remark 3.3(i) of [2], is false.

Remark 2. This theorem indicates that the Stackelberg solution is more beneficial to player 1 as should be expected in general than all the other Nash solutions under ECS and SIS. It is not difficult to see that the Nash solution under ECS considered in theorem 2 of [1] is actually a Stackelberg solution.

Remark 3. This theorem and Fig. 1 give a general description of the impact of ECS on J_1 which includes the result of theorem 2 of [1] as one particular impact out of uncountable ones.

Remark 4. The parameter d in (3) can be regarded as a measure of the strength with which player 1 communicates his control implicitly to player 2. It can be regarded also as an incentive mechanism in a leader–follower situation, e.g. if the leader cannot communicate his control value to the follower free from noise, then by designing $d = (1 + h_1^2 + h_2^2)b_1/h_2$ in (3) and playing Nash (ICS), the leader can expect the same performance as in a Stackelberg game where the follower has perfect knowledge of the leader's control value.

4. Comments

In this section we give comments concerning the impact of ECS on J_1. In the first part we explain part (i) of theorem 1, i.e. why a local minimum of J_1 among L is given by the Stackelberg solution instead of the SIS solution as claimed by Uchida. In the second part we explain part (ii) of theorem 1, i.e. why player 1 can do better in a continuous range of ζ under ECS than under SIS.

Since $J_i(\gamma_1, \gamma_2)$ is quadratic in γ_j, $i, j = 1, 2$, $J_i(\gamma_1, \gamma_2)$ is differentiable w.r.t. γ_j. Furthermore, γ_2 is

Stackelberg solution (6) to hold [4]. Since J_1 and J_2 are convex in γ_1 and γ_2, the first-order necessary condition is also a sufficient condition for ζ^* to be a Stackelberg solution.

It is remarkable that under ECS, although $J_1(\gamma_{1B}, \gamma_{2B})$ depends on the statistics of the observation noise, its ordering for different ζ is independent of the noise, as we can see from (20). Thus, in order to explain the ordering of $J_1(\gamma_{1B}, \gamma_{2B})$ for different ζ, one need consider only the deterministic game. In Section 2, if there is no observation noise in (2)–(4), then for any given u_1 the optimal value of u_2 minimizing J_2 is determined uniquely by

$$\gamma_2(x, u_1) = -(1 + q_2 b_2^2)^{-1} q_2 b_2 (b_1 u_1 + x) . \tag{29}$$

The locus of such points (u_1, u_2) given by (29) for all $u_1 \in R$ is called the reaction curve of player 2. The reaction curve of player 1 is similarly determined. Equicost contours of J_1 and J_2 and the reaction curves of both players are plotted in Fig. 3 for some particular values for the parameters of the game. The Nash solutions of Case B given by (10) now reduces to:

$$\gamma_{1B}(x) = -\{1 + q_1 b_1^2 + q_2 b_2^2 + \zeta(q_1 - p_{12} q_2) b_1 b_2\}^{-1} \{q_1 b_1 + \zeta(q_1 - p_{12} q_2) b_2\} x , \tag{30a}$$

$$\gamma_{2B}(x, u_1) = -(1 + q_2 b_2^2)^{-1} q_2 b_2 \{b_1 \gamma_{1B}(x) + x\} + \zeta(u_1 - \gamma_{1B}(x)) , \tag{30b}$$

for all $\zeta \in R$ such that (11) holds. Notice that at each solution point of (30), the value of u_2 given by (30b) is equal to that determined by the strategy

$$\gamma_2(x, u_1) = -(1 + q_2 b_2^2)^{-1} q_2 b_2 (b_1 u_1 + x) . \tag{31}$$

Equation (31) is the same as (29), which means that all the Nash solution pairs $(\gamma_{1B}, \gamma_{2B})_\zeta$ are on R_2, the reaction curve of player 2. Furthermore, since $\{u_1\}$ given by (30a) for all $\zeta \in R$ such that (11) holds, is the real line, we conclude that R_2 comprises all the linear Nash solutions of Case B. Point C in Fig. 3 represents $(\gamma_{1C}, \gamma_{2C}) = (\gamma_{1B}, \gamma_{2B})_{\zeta=0}$, the SIS solution where R_1 and R_2 intersect. Point S represents $(\gamma_{1S}, \gamma_{2S}) = (\gamma_{1B}, \gamma_{2B})_{\zeta=\zeta_S}$, the Stackelberg solution, where R_2 is tangent to the contour of J_1 [1]. Fig. 3 shows clearly that point S gives a global minimum of J_1 on R_2 and point C is by no means a local minimum of J_1 on R_2. All the points between C and S on R_2 yield lower cost of J_1 than point C. Finally,

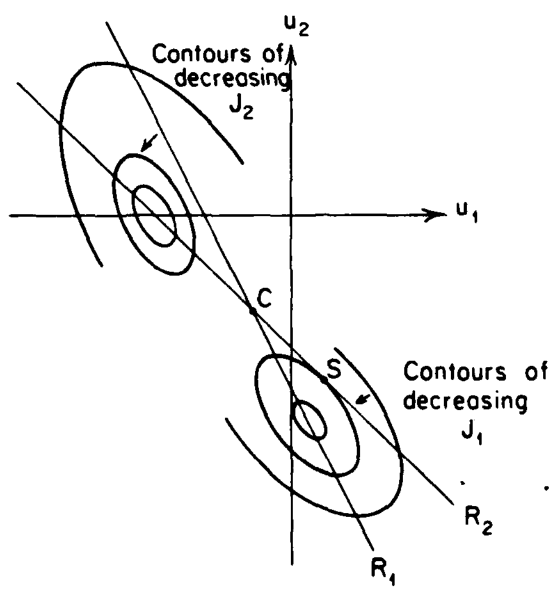


Fig. 3. Illustration of the impact of ECS on J_1. R_2: reaction curve of player 2; R_1: reaction curve of player 1.

3 Instanbul Workshop. June 1984.
submitted to SAM JOC
[This the revised version of previous title: "Adaptive Repeated Nash Games".]

ITERATIVE TECHNIQUES FOR THE NASH SOLUTION IN QUADRATIC GAMES WITH UNKNOWN PARAMETERS†

by

G. P. Papavassilopoulos
University of Southern California
Dept. of Electrical Engineering-Systems
Los Angeles, CA 90089-0781

ABSTRACT

We study adaptive schemes for repeated quadratic Nash games in a deterministic and a stochastic framework. The convergence of the schemes is demonstrated under certain conditions.

KEY WORDS

Nash equilibrium, Adaptive Games, Stochastic Approximation.

†This research was supported in part by the U. S. Air Force Office of Scientific Research under Grant AFOSR-82-0174 and by the USC Faculty Research and Innovation Fund.

1. INTRODUCTION

The object of this paper is the study of a static quadratic Nash game where the players do not have knowledge of the parameters involved in the description of the cost of their opponents and of their opponent's information. The game is played repeatedly and at each stage the players know the past actions of their opponents. The only dynamics involved are in the accumulation of the information on their opponent's previous actions; apart from this dynamic aspect, the problem considered is a repeated statis game. We examine both the deterministic and stochastic case, consider some adaptive schemes for updating the players decisions, and we show convergence to the optimal decisions (in the mean square sense and with probability one for the stochastic case), under some conditions. The scheme for the stochastic case is actually a stochastic approximation algorithm of the Robbins-Monro type.

The underlying motivation for the present paper is to study situations of conflict where the players do not know some of the parameters involved in the description of the others' cost functionals, or in the state equation. Such situations have been and are being studied for the single player – i.e., control problem – case and come under the name of Adaptive Control; the corresponding problems for situations of conflict, i.e., Adaptive Games, has received very little attention up to now. The problem studied here can be considered as a very simple type of adaptive game where the players adapt their decisions so as to converge in the limit to the solution of a static Nash game. It should be noted that the strategies exhibited in this paper do not constitute a Nash equilibrium pair for the construed dynamic – dynamic due to the dynamic information – game; but similarly, the adaptive control strategy in the self-tuning regulator problem [5], converges in the limit

to the optimal solution without being necessarily optimal at each stage. Adaptive games are important for several reasons. For example, when two players are involved in a situation of conflict, it is reasonable to assume that each player knows his own objective, but not that of his opponent; in addition, he might not know several of the parameters of the dynamic system which couples him with the other. In decentralized control, we think of decentralization as a scheme according to which each controller knows his own objective and information but not those of the others. If each controller knew the objectives of the others — as is implicitly assumed in many existing decentralized schemes — then the notion of decentralization is weakened. Although considerable progress has been achieved for the centralized controller, single objective adaptive control [4-6], the area of adaptive games is in its infancy. The only work that the author is familiar with in this area is [7] and [8]. In [7], adaptive schemes based on self-tuning for stochastic Nash and Stackelberg games are considered, where the players have the same information. (In the present paper the information of the players is different.) In [8] two adaptive schemes are studied for repeated Stackelberg games in a deterministic framework.

The structure of the paper is as follows. In Section 2 we consider the deterministic case and study three simple adaptive schemes. In Section 3 we consider an adaptive scheme for the stochastic case. The stochastic scheme is a Robbins-Monro type of stochastic approximation algorithm. Although several results exist for such algorithms, many of which can be used to provide convergence for the scheme considered here, the conditions of convergence that they would obtain for our scheme are more stringent than those that we prove here. In each section we provide several comments relating the results with previous work, expand on their meaning and provide appropriate motivation. Finally, we have a conclusions section.

2. DETERMINISTIC CASE

Let J_1, J_2: $R^{m_1} \times R^{m_2} \to R$ be two functions defined by:

$$J_i(u_1,u_2) = \tfrac{1}{2}\, u_i'u_i + u_i'R_iu_j + u_i'c_i, \quad i \neq j, \quad i,j = 1,2 \qquad (1)$$

where $u_i \in R^{m_i}$, R_1, R_2 are real constant matrices and c_1, c_2 are real constant vectors of appropriate dimensions. A pair (u_1^*,u_2^*) is a Nash equilibrium if it satisfies ([1],[2]):

$$J_1(u_1^*,u_2^*) \leq J_1(u_1,u_2^*), \quad \forall\, u_1 \in R^{m_1} \qquad (2)$$

$$J_2(u_1^*,u_2^*) \leq J_2(u_1^*,u_2), \quad \forall\, u_2 \in R^{m_2} \qquad (3)$$

or equivalently if

$$R \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix} + c = 0, \quad R = \begin{bmatrix} 1 & R_1 \\ R_2 & 1 \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \qquad (4)$$

J_i and u_i are the cost and the decision of player i.

Let us assume that player i knows R_i and c_i, but not R_j and c_j $(j \neq i)$; then he cannot solve (4) for u_i^*. Consider also that this game is played repeatedly at times $t = 1,2,3,\ldots$, that at time t, player i knows $I_t^i = \{u_{1,1},\ldots,u_{1,t-1}, u_{2,1},\ldots,u_{2,t-1}\}$ and plays u_{it} which is chosen as a function of I_t^i, i.e.,

$$u_{it} = F_i(I_t^i, t), \quad i = 1,2, \quad t = 2,3,\ldots \qquad (5)$$

The question is: For what F_1, F_2 the recursion (5) will converge to a solution of (4). Let us now examine three possible choices of F_1, F_2.

Case 1

$$F_i(I_t^i,t) = -R_i u_{j,t-1} - c_i, \quad i=1,2, \quad i \neq j \tag{6}$$

The meaning of (6) is that player 1 minimizes $J_1(u_1,u_{2,t-1})$, i.e., he reacts only to the last announced decision of player 2. Recursion (5) assumes the form:

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} - \left(R \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + c \right), \quad t \geq 2 \tag{7}$$

Recursion (7) will converge to a solution of (4) for any initial condition $(u_{1,1},u_{2,1})$ if and only if all the eigenvalues of the matrix R lie within the open disc of radius 1 centered at the point 1 in the complex plane, i.e.,

$$|\lambda(R) - 1| < 1 \tag{8}$$

((8) is equivalent to: $|\lambda(R_1R_2)| < 1$.) Condition (8) also guarantees that (4) has a unique solution.

Case 2

$$F_i(I_t^i,t) = -R_i[u_{j,t-1} + \theta u_{j,t-2} + \dots + \theta^{t-2}u_{j,1}]\,\frac{1-\theta}{1-\theta^{t-1}} - c_i \tag{9}$$

$$1 > \theta \geq 0, \quad i=1,2, \quad i \neq j$$

The meaning of (9) is that player 1 minimizes J_1 with respect of u_1, with u_2 fixed to a value that is a weighted average of $u_{2,t-1},\ldots,u_{2,1}$ where more weight is put on the recent values of u_2. We assume that both players use the same θ. Recursion (9) can be written equivalently:

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} - \frac{1-\theta}{1-\theta^{t-1}} \left(R \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + c \right), \ t \geq 2 \qquad (10)$$

Recursion (10) will converge to a solution of (4) for any initial condition $(u_{1,1}, u_{2,1})$ if and only if all the eigenvalues of the matrix R lie within the open disc of radius $(1-\theta)^{-1}$ centered at the point $(1-\theta)^{-1}$ in the complex plane, i.e.,

$$\left| \lambda(R) - \frac{1}{1-\theta} \right| < \frac{1}{1-\theta} \qquad (11)$$

Condition (11) also guarantees that (4) has a unique solution. (Notice that as $t \to +\infty$, $\theta^{t-1} \to 0$ and thus $(1-\theta)R$ in (10) assumes the role of R in (7).) Obviously, for $\theta = 0$, (11) reduces to (8) and (10) to (7).

Case 3

$$F_i(I_t^i, t) = -R_i[u_{j,t-1} + u_{j,t-2} + \ldots + u_{j,1}] \frac{1}{t-1} - c_i \qquad (12)$$

$$i = 1,2, \quad i \neq j.$$

The meaning of (12) is that player 1 minimizes J_1 with respect to u_1, with u_2 fixed to the arithmetic mean of $u_{2,t-1},\ldots,u_{2,1}$. Recursion (12) can be

written equivalently:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} - \frac{1}{t-1}\left(R \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + c\right), \quad t \geq 2 \qquad (13)$$

Recursion (13) will converge to a solution of (4), for any initial condition $(u_{1,1}, u_{2,1})$ if and only if all the eigenvalues of R has positive real parts, i.e.,

$$\text{Re}\ \lambda(R) > 0 \qquad (14)$$

(For proof see Appendix A, Lemma A3.) Condition (14) also guarantees that (4) has a unique solution. Notice that as $\theta \to 1$, (11) reduces to (14).

Remark 1 Obviously (8) $\Rightarrow$ (11) $\Rightarrow$ (14). If (8) holds, (7) converges faster than (10) and if (11) holds, (10) converges faster than (13).

Remark 2 In all three cases we assumed that both players use the same scheme. Nonetheless, it might happen that they use different ones. It is easy to verify that if player 1 uses scheme 1 and player 2 uses scheme 2, the region of convergence is larger than if both were using scheme 1 and worse than if both were using scheme 2. Similar results hold for the other combinations.

Remark 3 If we consider (10) with $\theta > 1$, i.e., more weight is assigned to the old measurements, the scheme will not converge. This can be easily verified by considering the scalar version of (10) with $c = 0$:

$$u_t = u_{t-1}\left(1 - r\,\frac{1-\mu}{\mu}\,\frac{\mu^{t-1}}{1-\mu^{t-1}}\right),\ \mu = \frac{1}{\theta}$$

which for $t \to +\infty$ behaves like

$$y_t = y_{t-1}\left(1 - r\,\frac{1-\mu}{\mu}\,\mu^{t-1}\right)$$

(since $0 < \mu < 1$) and is easily seen to fail to converge.

Remark 4 (8), (11) and (14) can be expressed equivalently in terms of the eigenvalues of R_1R_2.

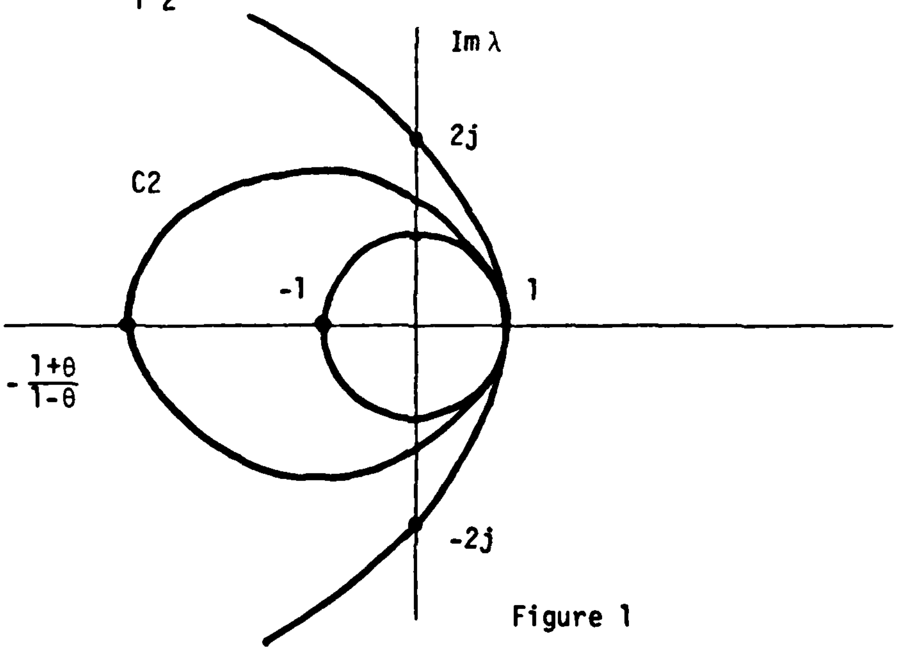


Figure 1

(8) corresponds to $|\lambda(R_1R_2)| < 1$, i.e., inside the unit disc, (11) corresponds to

$$(1-\theta)\ |\lambda| \pm 2\theta \cos\frac{\varphi}{2}\,|\lambda|^{\frac{1}{2}} - (1+\theta) < 0$$

$$\lambda(R_1R_2) = |\lambda| e^{j\varphi}$$

i.e., inside the curve C2 of Fig. 1. (14) corresponds to eigenvalues of R_1R_2 being inside the parabola defined by

$$\mathrm{Re}\,\lambda + \tfrac{1}{4}(\mathrm{Im}\lambda)^2 < 1, \quad \lambda = \lambda(R_1R_2) .$$

Remark 5 If (8) (or equivalently $|\lambda(R_1R_2)| < 1$) holds, the solution of (4) is called in game theory a stable equilibrium, and the game is called stable [1]. The reason is that if player i deviates from u_i^*, then player $j(j \neq i)$ responds according to scheme (6) and to that player i responds according to scheme (6) and so on and eventually they both converge back to (u_1^*, u_2^*). Obviously the notion of stable equilibrium depends on the reaction scheme that the players employ. If schemes (9) or (12) are used as reaction schemes, we have an enlarged class of stable games.

Remark 6 Since the scheme of case 3 (12) has the best convergence region out of the three schemes, in the next section we will deal with the stochastic analogue of (12).

Remark 7 All three schemes considered, can actually be viewed as schemes for solving $Ru + c = 0$ (see (4)), by using an iteration of the form:

$$u_{n+1} = u_n - D_n[Ru_n + c] \tag{15}$$

where D_n has to have the structure

$$D_n = \begin{bmatrix} D_n^1 & 0 \\ 0 & D_n^2 \end{bmatrix} .$$

(Iterative solutions of linear equations is a vast subject, see for e.g. [16].) Scheme (13) employed: $D_n^i = \frac{1}{n} I$. We can create new schemes which converge under weaker conditions than (14) by allowing $D_n^i = \frac{1}{n} D^i$ where D^1, D^2 are properly chosen constant matrices. For example, if R_1, R_2 are scalars, (14) is equivalent to $1 > r_1 r_2$; but if we use $D_n^i = \frac{1}{n} d_i$ in (15), the convergence condition becomes

$$\mathrm{Re}\lambda\left(\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}\begin{bmatrix} 1 & r_1 \\ r_2 & 1 \end{bmatrix}\right) > 0$$

which is equivalent to:

$$d_1 + d_2 > 0$$

$$d_1 d_2 (1 - r_1 r_2) > 0$$

and can always be satisfied for some d_1, d_2 as long as $1 \neq r_1 r_2$. Notice, that $1 \neq r_1 r_2$ is the necessary and sufficient condition for solvability of (4) for any c.

Remark 8 Another way of going about the problem of this section is to consider that at each stage, each player uses a certain scheme to estimate the R and C of his opponent and then calculates his action by solving (4_1) wherein he employs the estimates of the R and c of his opponent. In such a scheme, each player should know at each stage not only the previous actions of his opponent – as in our scheme – but also the rationale according to which his opponent calculates his actions. This is necessary in order just to estimate his opponent's parameters at each stage. Nonetheless, such an additional

knowledge can be permitted and the convergence of the resulting scheme studied. Finally, it should be noted that the problem considered here and the schemes proposed, besides having their own merit, provide a certain motivation for the schemes considered for the stochastic case of the next section.

3. THE STOCHASTIC CASE

Let x be a Gaussian random vector in R^n with zero mean and unit covariance matrix. Let

$$y_i = C_i x, \quad i = 1,2 \tag{16}$$

represent the measurements of the two players, where C_1, C_2 are fixed real matrices of dimensions $n_1 \times n$, $n_2 \times n$ respectively. Let Γ_i be the set of all measurable $\gamma_i : R^{n_i} \to R^{m_i}$ functions with $E[\gamma_i(y_i)'\gamma_i(y_i)] < +\infty$. Set $u_i = \gamma_i(y_i)$ and let

$$J_i(\gamma_1,\gamma_2) = E[\tfrac{1}{2}\, u_i'u_i + u_i'R_i u_j + u_i'S_i x], \; i \neq j, \quad i,j = 1,2 \tag{17}$$

represent the costs of the two players. R_1, R_2, S_1, S_2 are fixed real matrices of appropriate dimensions. A pair $(\gamma_1^*,\gamma_2^*) \in \Gamma_1 \times \Gamma_2$ is called a Nash equilibrium if it satisfies

$$J_1(\gamma_1^*,\gamma_2^*) \leq J_1(\gamma_1,\gamma_2^*) \qquad \forall\, \gamma_1 \in \Gamma_1 \tag{18}$$

$$J_2(\gamma_1^*,\gamma_2^*) \leq J_2(\gamma_1^*,\gamma_2) \qquad \forall\, \gamma_2 \in \Gamma_2$$

For background concerning the formulation of the stochastic Nash game see [18]. (18) is equivalent to (see [2,3]):

$$\gamma_1^*(y_1) + R_1\, E[\gamma_2^*(y_2)|y_1] + S_1 E\,[x|y_1] = 0 \tag{19a}$$

$$\gamma_2^*(y_2) + R_2\, E[\gamma_1^*(y_1)|y_2] + S_2 E\,[x|y_2] = 0 \tag{19b}$$

It is known (see [3]) that if no eigenvalue of R_1R_2 equals the inverse of any arbitrary but finite product of powers of the squares of the canonical correlation coefficients of y_1, y_2 (i.e., of $\sigma_1, \sigma_2, \ldots$), then (19) has a unique solution which has to be linear in the information. The set of values where the eigenvalues of R_1R_2 should not lie is a countable isolated set of points in $[1,+\infty)$ and thus it is generically true that (19) admits a unique solution which has to be linear in the information. We can assume without loss of generality (see Lemma 1 [3]) that

$$n_1 \leq n_2 \quad C_1C_1' = I_{n_1 \times n_1}, \; C_2C_2' = I_{n_2 \times n_2}, \; C_1C_2' = \begin{bmatrix} \sigma_1 & & 0 & \vdots & \\ & \sigma_2 & & \vdots & 0 \\ 0 & & \ddots \; \sigma_{n_1} & \vdots & \end{bmatrix}_{n_1 \times n_2} \qquad (20)$$

$$1 \geq \sigma_1 \geq \sigma_2 \geq \sigma_{n_1} \geq 0$$

and then $\gamma_i^*(y_i) = L_iy_i$ where L_1, L_2 are the solutions to the system:

$$L_1 + R_1L_2C_2C_1' + S_1C_1' = 0 \qquad (21)$$

$$L_2 + R_2L_1C_1C_2' + S_2C_2' = 0$$

Let us assume that player i knows R_i, S_i, C_i, but not R_j, S_j, C_j, $i \neq j$; then he cannot solve (21) for L_i. Consider also that this game is played repeatedly at times $t = 1,2,3,\ldots$, that at time t player i knows

$$I_t^i = \{u_{1,1},\ldots,u_{1,t-1},u_{2,1},\ldots,u_{2,t-1},y_{i,1},\ldots,y_{i,t-1}\} \quad (22)$$

where y_{it} is the measurement of player i at time t. We assume that

$$y_{it} = C_i x_t \tag{23}$$

where the x_t's are independent Gaussian vectors with zero mean and unit covariance. At time t, player 1 employs the following scheme for finding u_{1t}:

$$u_{1t} + R_1\left(\frac{1}{t-1}\sum_{k=1}^{t-1} u_{2,k}y'_{1,k}\right)y_{1t} + S_1C'_1y_{1t} = 0 \tag{24}$$

A justification of this scheme is the following: at time t player 1 has to solve (19a) for u_{1t} and thus he has to calculate $E[u_{2,t}|y_{1t}]$, $E[x|y_{1t}]$. If u_{2t} is linear in y_{2t}, then u_{2t}, y_{1t} are jointly gaussian and thus

$$E[u_{2,t}|y_{1t}] = E[u_{2t}y'_{1t}]\,(E[y_{1t}y'_{1t}])^{-1}y_{1,t} \tag{25}$$

Player 1 approximates $E[u_{2t}y'_{1t}]$ by $\frac{1}{t-1}\sum_{k=1}^{t-1}(u_{2,k}y^T_{1k})$; a motivation for this approximation is the following: If player 1 knew all the parameters of (16), (17), he would then solve equation (19) at stage t, employing (23); due to the independence of the x_t's, $\frac{1}{t-1}\sum_{k=1}^{t-1}(u_{2k}y^T_{1k})$ would provide a reasonable approximation of $E[u_{2t}|y_{1t}]$, since u_{2k} would be independent of $u_{2\ell}, y_{1\ell}$, $\ell \neq k$. By overlooking the lack of independence of u_{2k} on $u_{2\ell}, y_{1\ell}$, $\ell \neq k$, he still employs the above approximation, hoping that things will work out. The convergence results of Theorem 1' and 2' provide a posterior justification for the reasonableness of this approximation.

By our assumption (20) $E[y_{1t}y'_{1t}] = I$ and $E[x_t|y_{1t}] = S_1C'_1y_{1t}$. (24) yields that u_{1t} is linear in y_{1t}, i.e., $u_{1t} = L_{1t}y_{1t}$ where L_{1t} satisfies

$$L_{1t} + R_1 \left[\frac{1}{t-1} \sum_{k=1}^{t-1} u_{2k} y_{1k}'\right] + S_1 C_1' = 0 \tag{26}$$

A similar equation is satisfied by L_{2t}, if we consider that u_{2t} is calculated by an equation corresponding to (24) and $u_{2t} = L_{2t} y_{2t}$. The equations for L_{1t}, L_{2t} can be written recursively as:

$$L_{1t} = L_{1,t-1} - \frac{1}{t-1} [L_{1,t-1} + R_1 L_{2,t-1} y_{2,t-1} y_{1,t-1}' + S_1 C_1'] \tag{27a}$$

$$L_{2t} = L_{2,t-1} - \frac{1}{t-1} [L_{2,t-1} + R_2 L_{1,t-1} y_{1,t-1} y_{2,t-1}' + S_2 C_2'] \tag{27b}$$

(27) is the recursion that we intend to study and show that under some conditions converges to the solution of (21) in the q.m. sense and w.p.l. The initial condition L_{11}, L_{21} of (27) is taken to be an arbitrary pair of real constant matrices and we are interested in convergence for any initial condition. (27) defines a Markovian stochastic process (L_{1t}, L_{2t}) and is obviously a stochastic approximation algorithm of the Robbins-Monro type [9] for solving (21). (27) is the stochastic analogue of the scheme of case 3 of the deterministic case.

Let us now study the convergence of (27). Let us call ℓ_{it}, m_{it}, c_i, d_i the i-th columns of L_{1t}, L_{2t}, S_1C_1', S_2C_2' respectively, i.e.,

$$\begin{aligned} &L_{1t} = [\ell_{1t}, \ldots, \ell_{n_1 t}], \quad L_{2t} = [m_{1t}, \ldots, m_{n_2 t}] \\ &S_1 C_1' = [c_1, \ldots c_{n_1}], \quad S_2 C_2' = [d_2, \ldots, d_{n_2}] \end{aligned} \tag{28}$$

Let

$$\bar{\ell}_{it} = E[\ell_{it}], \qquad \bar{m}_{it} = E[m_{it}] \tag{29}$$

Using (20) and the fact that L_{1t} depends on $y_{11},\ldots,y_{1t-1},y_{21},\ldots,y_{21},\ldots,y_{2,t-1}$, we obtain from (27):

$$\bar{\ell}_{it} = \bar{\ell}_{it-1} - \frac{1}{t-1}[\bar{\ell}_{i,t-1} + \sigma_i R_1 \bar{m}_{i,t-1} + c_i] \tag{30a}$$

$$\bar{m}_{it} = \bar{m}_{it-1} - \frac{1}{t-1}[\bar{m}_{i,t-1} + \sigma_i R_2 \bar{\ell}_{i,t-1} + d_i] \tag{30b}$$

$$i = 1,\ldots,n_1$$

and

$$\bar{m}_{it} = \bar{m}_{i,t-1} - \frac{1}{t-1}[\bar{m}_{i,t-1} + d_i]$$

$$i = n_1+1,\ldots,n_2 \tag{30c}$$

Recursion (30c) converges for any initial condition (see Lemma A3). (30a) can be written as

$$\begin{bmatrix} \bar{\ell}_{it} \\ \bar{m}_{it} \end{bmatrix} = \begin{bmatrix} \bar{\ell}_{i,t-1} \\ \bar{m}_{i,t-1} \end{bmatrix} - \frac{1}{t-1}\left(\begin{bmatrix} I & \sigma_i R_1 \\ \sigma_i R_2 & I \end{bmatrix}\begin{bmatrix} \bar{\ell}_{i,t-1} \\ \bar{m}_{i,t-1} \end{bmatrix} + \begin{bmatrix} c_i \\ d_i \end{bmatrix}\right) \tag{31}$$

and using Lemma A3 yields that (31) converges for any initial condition if and only if

$$\operatorname{Re}\lambda\left(\begin{bmatrix} I & \sigma_i R_1 \\ \sigma_i R_2 & I \end{bmatrix}\right) > 0 \tag{32}$$

It is easy to see that if (32) holds for σ_1 then it holds for any σ_i, $0 \le \sigma_i \le \sigma_1$. We thus have proven:

<u>Theorem 1'</u> The means of L_{1t}, L_{2t} as defined by the recursion (27) converge to a solution of (21) for any initial condition, if and only if

$$\operatorname{Re}\lambda\left(\begin{bmatrix} I & \sigma_1 R_1 \\ \sigma_1 R_2 & I \end{bmatrix}\right) > 0 \tag{33}$$

It is easy to see that if (33) holds then (21) has a unique solution. If we want (27) to converge to a solution of (21) not only for any initial condition, but also for any pair of measurements, i.e., any C_1, C_2, we have to consider $\sigma_1 = 1$ in (33) which is exactly the condition for convergence of case 3 of the determinstic case.

Next we will show that L_{1t}, L_{2t} converge to a solution of (21) in the mean square sense, under condition (33). For simplicity and w.l.o.g. we will assume $S_1C_1' = 0$, $S_2C_2' = 0$. We can write (27) component wise in terms of ℓ_{it}, m_{it} and then form the products $\ell_{it}\ell_{jt}'$, $i,j = 1,\ldots,n_1$, $m_{it}m_{jt}'$, $i,j = 1,\ldots, n_2$ and $\ell_{it}m_{jt}'$ $i = 1,\ldots,n_1$, $j = 1,\ldots,n_2$. These products satisfy recursions that can be easily calculated, and taking expectations of which result in a recursion which gives the evolution of $E(\ell_{it}\ell_{jt}')$, $E(m_{it}m_{jt}')$, $E(\ell_{it}m_{jt}')$ in terms of $E(\ell_{i,t-1}\ell_{j,t-1}')$, $E(m_{i,t-1},m_{j,t-1}')$ $E(\ell_{i,t-1}m_{j,t-1}')$. Before writing down

this recursion we introduce some notation:

$$\Lambda_{ij}^{t} = E[\ell_{it}\ell_{jt}'], \qquad i,j = 1,\ldots,n_1 \tag{34a}$$

$$M_{ij}^{t} = E[m_{it}m_{jt}'], \qquad i,j = 1,\ldots,n_2 \tag{34b}$$

$$K_{ij}^{t} = E[\ell_{it}m_{jt}'], \qquad i = 1,\ldots,n_1 \qquad j = 1,\ldots,n_2 \tag{34c}$$

$$N_t = \left[\begin{array}{ccc:ccc}
\Lambda_{11}^{t} & \cdots & \Lambda_{1n_1}^{t} & K_{1,1}^{t} & \cdots & K_{1,n_2}^{t} \\
\vdots & & \vdots & \vdots & & \vdots \\
\Lambda_{n_1 1}^{t} & \cdots & \Lambda_{n_1 n_1}^{t} & K_{n_1,1} & \cdots & K_{n_1,n_2}^{t} \\
\hdashline
(K_{11}^{t})' & \cdots & (K_{n_1,1}^{t})' & M_{11}^{t} & \cdots & M_{1,n_2}^{t} \\
\vdots & & \vdots & \vdots & & \vdots \\
(K_{1,n_2}^{t})' & \cdots & (K_{n_1 n_2}^{t})' & M_{n_2,1} & \cdots & M_{n_2,n_2}^{t}
\end{array}\right] \tag{35}$$

$$Q = \left[\begin{array}{ccc:cccc}
I & & 0 & \sigma_1 R_1 & 0 & 0 & 0 \\
 & \ddots & & & \sigma_2 R_1 \ddots & \vdots & \\
0 & & I & 0 & \sigma_{n_1} R_1 & 0 \cdots & 0 \\
\hdashline
\sigma_1 R_2 & \ddots & 0 & I & & & \\
 & 0 & \sigma_{n_1} R_2 & & \ddots & & 0 \\
0 & \cdots & 0 & 0 & & \ddots & \\
\vdots & & \vdots & & & & \\
0 & \cdots & 0 & & & & I
\end{array}\right] \tag{36}$$

Then N_t satisfies:

$$N_t = N_{t-1} - \frac{1}{t-1}[N_{t-1}Q' + QN_{t-1}] + \frac{1}{(t-1)^2}\mathscr{L}(N_{t-1}) \ . \tag{37}$$

where $\mathscr{L}(\cdot)$ denotes a linear time invariant function of its argument. (For details of this derivation, see Appendix B.)

Using Lemma A4 we conclude that N_t goes to zero for any initial condition if and only if the matrix Q has eigenvalues with positive real parts which is easily seen to be equivalent to (33). We thus have proven

Theorem 2' L_{1t}, L_{2t} as defined by recursion (27) converge to a solution of (21) for any initial condition, in the mean square sense, if and only if (33) holds.

Next, we will show that (L_{1t}, L_{2t}) converges under (33) for any initial condition to the solution of (21) with probability 1 (i.e., a.s. convergence). We again assume for simplicity and w.l.o.g. that $S_1C_1' = 0$, $S_2C_2' = 0$. We will use the theorem in paragraph 3 of [11] (or Lemma 3.5 of [13]) which we restate here and which is an easy consequence of the martingale convergence theorem of Doob.

Lemma 1 Let $\{V_t\}$ be a sequence of random variables such that $E(V_1)$ exists. Let A be a real number and suppose $V_t \geq A$. Furthermore, assume that $\sum_{t=1}^{\infty} E(E[V_{t+1} - V_t | V_1, \dots, V_t]^+)$ converges. Then the sequence $\{V_t\}$ converges with probability 1.

(Recall that if x is a random variable: $x^+ = \frac{1}{2}(|x| + x)$.) Let $x_t = (\ell_{1t}', \dots, \ell_{n_1,t}', m_{1,t}', \dots, m_{n_2,t}')'$. We will prove that x_t converges to 0 w.p.1 or equivalently that $V_t = \|x_t\|^2$ does. Let A = 0. From (27) we can easily

obtain (see Appendix C)

$$|E\,[V_{t+1} - V_t | V_1, \ldots, V_t]| \leq \frac{\alpha}{t} V_t$$

for some positive number α and thus

$$E[V_{t+1} - V_t | V_1, \ldots, V_t]^+ \leq \frac{\alpha}{t} V_t$$

In order to fulfill the assumption of Lemma 1, it suffices to show that

$$\sum_{t=1}^{\infty} \frac{\alpha}{t} E(V_t) < +\infty \tag{38}$$

It holds

$$E[V_t] = \operatorname{tr} N_t$$

and thus it suffices to show that

$$\sum_{t=1}^{\infty} \frac{\operatorname{tr} N_t}{t} < +\infty \tag{39}$$

From (37) we obtain

$$\begin{aligned} N_{t+1} = N_1 - Q \left[\sum_{t=1}^{t} \frac{N_k}{k} \right] - \left[\sum_{k=1}^{t} \frac{N_k}{k} \right] Q' + \\ + \mathcal{L} \left(\sum_{k=1}^{t} \frac{N_k}{k^2} \right) \end{aligned} \tag{40}$$

If we assume that Q has eigenvalues with positive real parts (40) can be solved for $\sum_{k=1}^{t} \frac{N_k}{k}$ to yield

$$\sum_{k=1}^{t} \frac{N_k}{k} = \mathcal{L}'\left(N_{t+1}, N_1, \sum_{k=1}^{t} \frac{N_k}{k^2}\right)$$

Since N_k converges, it is bounded and so is $\sum_{k=1}^{t} \frac{N_k}{k^2}$. Thus $\sum_{k=1}^{t} \frac{N_k}{k}$ is uniformly bounded and thus (39) and (38) are bounded. We thus conclude that $\|x_t\|^2 = V_t$ converges with probability 1. $\|x_t\|^2$ converges to 0 in the mean square sense by Theorem 2' and thus in probability and thus it has a subsequence converging to zero with probability one ([17], Thm. 2, 5, 3, p. 93). Since we just showed that $\|x_t\|^2$ converges with probability one, this limit has to be zero. Let us now summarize the results of this section in a Theorem.

Theorem: L_{1t}, L_{2t} as defined by recursion (27) converge to a solution of (21) for any initial condition, in the mean square sense and with probability one if and only if

$$\operatorname{Re}\lambda\left(\begin{bmatrix} I & \sigma_1 R_1 \\ \sigma_1 R_2 & I \end{bmatrix}\right) > 0$$

(Under this condition (21) admits a unique solution.)

Remark 1 N_t, (37), goes to zero but it does not have to converge monotonically.

Remark 2 One can construct the stochastic analogues of the deterministic schemes of cases 1 and 2, if a different — appropriate — approximation is used

for $E[u_{2t}|y_{1t}]$ in (25). A little reflection, though, will persuade the reader that these schemes will converge under conditions more stringent than (33).

Remark 3 For a repeated Stackelberg game one can consider schemes similar to those considered here, if one assumes that the Leader does not know the parameters involved in the Follower's cost. An idea of this sort was recently studied in a deterministic framework in [8].

Remark 4 It should be clear from (30) and (37) that the rate of convergence of the means and the covariances of ℓ_{it}, m_{it} depend on the eigenvalues of the matrices in (32) for $\sigma_i = 1, \sigma_1, \ldots, \sigma_{n_1}$, or equivalently of Q. Actually, a recursion of the form (A1) with $\bar{\lambda} = \mathrm{Re}(\lambda) > 0$ goes to zero like $(n^{\bar{\lambda}})^{-1}$ (see [12]). Thus if λ_m denotes the real part of the eigenvalues of Q, $m = 1, \ldots, n_1 + n_2$ and $\bar{\lambda} = \min \mathrm{Re}(\lambda_m)$ the mean converges no slower than $(t^{\bar{\lambda}})^{-1}$, the covariances no slower than $(t^{2\bar{\lambda}})^{-1}$, the third moments no slower than $(t^{3\bar{\lambda}})^{-1}$ and so on. Thus if one were to consider whether $t^{\theta}[L_{1t}, L_{2t}]$ converges weakly to a gaussian random variable as $t \to \infty$, θ should be chosen equal to $\bar{\lambda}$ so that the second moments converge to a nonzero constant, but then automatically all the moments will also do so. Thus in general one cannot have asymptotic normality of $n^{\theta}[L_{1t}, L_{2t}]$ for some $\theta > 0$. As a matter of fact, Theorem (1) of [12] cannot be applied since it's assumption (A4) fails for the stochastic approximation algorithm (27), considered here, as should be expected from the above remarks. Finally, it should be pointed out that the fact that the rate of convergence of the algorithm is given by $t^{-\bar{\lambda}}$ and $t^{-2\bar{\lambda}}$ for the first and second moments, is a useful fact when implementing it, in deciding when to stop, what is the probability of error when stopping in a finite number of iterations, etc.

Remark 5 Stochastic approximation has been an object of intensive study (see [9-15]). Several of the results available can be used to prove convergence of the iteration (27) but they demand conditions stronger than (33), or they are not applicable to it. For example, in [9] it is required that in the scheme $x_{n+1} = x_n - \frac{1}{n} y_n$, y_n is uniformly bounded. Assumptions III and IV of [10] do not hold for (27). In proving asymptotic normality [12], he uses Assumption (A4) which does not hold for (27). Assumptions A5, A5' of [11] do not hold for our scheme. Lemma 3.1 and Theorem 4.3 of [13] can be applied to (27) but result in more stringent conditions than (33). The convergence analysis of [15] demands boundness of the second term in (27) which is not applicable to our case. Assumption iii in Problem 1, p. 92 of [14] does not hold for (27).

4. CONCLUSIONS

There are several directions in which this research can be continued. One of them is the corresponding problem for the Stackelberg game (see Remark 3 in Section 3). The dynamic case where the players are also coupled through the evolution of a discrete time equation is obviously important and useful. We hope that the analysis presented here will be helpful in such further research.

APPENDIX A

Lemma A1 Consider the scalar recursion

$$x_{n+1} = \left(1 - \frac{\lambda}{n}\right) x_n, \quad n = 1,2,3,\ldots \tag{A1}$$

where λ and x_1 are complex numbers. Then $x_n \to 0$ for any x_1 if and only if $\mathrm{Re}(\lambda) > 0$. (If we set $t_n = 1 + \ldots + \frac{1}{n}$, we see that (A1) is a discrete approximation of $\dot{x} = -\lambda x$ and thus $\mathrm{Re}(\lambda) > 0$ is expected in order to have asymptotic stability of (A1).)

Lemma A2 Consider the scalar recursion

$$x_{n+1} = \left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n^2}\right)\right)x_n, \quad n = 1,2,3,\ldots$$

where x and x_1 are complex numbers. Then $x_n \to 0$ for any x_1 if and only if $\mathrm{Re}(\lambda) > 0$.

Proof

It is an immediate consequence of Lemma A1 since $\frac{\lambda}{n}$ dominates $O(\frac{1}{n^2})$. □

Lemma A3 Consider the recursion

$$x_{n+1} = \left(I - \frac{1}{n} A + O\left(\frac{1}{n^2}\right)\right)x_n, \quad n = 1,2,3,\ldots \tag{A3}$$

where A is a real square matrix and x_1 is a vector. Then $x_n \to 0$ for any x_1 if and only if $\mathrm{Re}\lambda(A) > 0$.

Proof We bring A to it's Jordan form and apply Lemma A2. It is helpful to notice that if P is a real symmetric matrix

$$x'_{n+1}Px_{n+1} = x'_nPx_n - \frac{1}{n} x'_n [PA + A'P] x_n + x'_n O(\frac{1}{n^2})x_n$$

and thus if A has $\mathrm{Re}\lambda(A) > 0$, we can find a positive definite P so that $A'P + PA > 0$. Therefore if n is sufficiently large

$$\frac{1}{n} x'_n[PA + A'P] x_n > x'_n O(\frac{1}{n^2}) x_n$$

and thus $x'_{n+1}Px_n < x'_nPx_n$ and consequently x_n is bounded. This justifies the fact that the $\frac{1}{n}$ term dominates in (A3). □

Lemma A4 Consider the recursion

$$N_{t+1} = N_t - \frac{1}{t} [N_tQ' + QN_t] + \frac{1}{t^2} \mathcal{L} (N_t), \quad t = 1,2,... \tag{A4}$$

where N_t, Q are square matrices. $N_t \to 0$ for any initial condition if and only if $\mathrm{Re}\ \lambda(Q) > 0$.

Proof Let x_t be the vector composed of the columns of N_t. We can write the recursion equivalently as

$$x_{t+1} = x_t - \frac{1}{t} Ax_t + \frac{1}{t^2} \mathcal{L}(x_t)$$

It can be checked that $\mathrm{Re}\ \lambda(A) > 0$ if and only if $\mathrm{Re}\ \lambda(Q) > 0$ and thus Lemma A3 can be applied. □

It should be pointed out that if x_n evolves as in A1, and λ is real, x_n behaves like $n^{-\lambda}$ (see [12], eq. 2.3). If λ is complex, then (A2) implies that $|x_n|^2$ behaves like n^{-2a} and thus $|x_n|$ behaves like n^{-a}, i.e., $n^{-\mathrm{Re}\,\lambda}$. Consequently x_{n+1} in (A3) behaves like $n^{-\bar{\lambda}}$, where $\bar{\lambda} = \min \mathrm{Re}\lambda\,(A)$ and N_t in (A4) behaves like $t^{-2\hat{\lambda}}$ where $\hat{\lambda} = \min \mathrm{Re}\,\lambda\,(Q)$.

APPENDIX B

Let ℓ_{it}, m_{it}, c_i, d_i be as in (28). For convenience, let

$$y_{i,t-1} = \begin{bmatrix} y_1 \\ \vdots \\ y_{n_1} \end{bmatrix}, \quad y_{2,t-1} = \begin{bmatrix} z_1 \\ \vdots \\ z_{n_2} \end{bmatrix} \tag{B1}$$

(27) can be written as

$$\left.\begin{aligned} \ell_{it} = \ell_{i,t-1} - \frac{1}{t-1}\left[\ell_{i,t-1} + y_i R_1 \sum_{j=1}^{n_2} z_j m_{j,t-1} + c_i\right] \\ i = 1,\ldots,n_1 \end{aligned}\right\} \tag{B2}$$

$$\left.\begin{aligned} m_{it} = m_{i,t-1} - \frac{1}{t-1}\left[m_{i,t-1} + z_i R_2 \sum_{j=1}^{n_1} y_j \ell_{j,t-1} + d_i\right] \\ i = 1,\ldots,n_2 \end{aligned}\right\} \tag{B3}$$

For convenience, let us drop the subscript t-1 from $\ell_{i,t-1}$, $m_{i,t-1}$. From (B2), (B3), we obtain:

$$\ell_{it}\ell'_{jt} = \ell_i\ell'_j - \frac{1}{t-1}\left[2\ell_i\ell'_j + y_j \sum_{\ell=1}^{n_2} z_\ell \ell_i m'_\ell R'_1 + y_i R_1 \sum_{k=1}^{n_2} z_k m_k \ell'_j +\right.$$

$$\left. + \ell_i c'_j + c_i \ell'_j\right] + \frac{1}{(t-1)^2}\left[\ell_i\ell'_j + y_j \sum_{\ell=1}^{n_2} z_\ell \ell_i m'_\ell R'_1 +\right.$$

$$+ y_i R_1 \sum_{k=1}^{n_2} z_k m_k \ell'_j + y_i y_j R_1 \sum_{k,\ell=1}^{n_2} z_k z_\ell m_k m'_\ell R'_1 +$$

$$+ y_i R_1 \sum_{k=1}^{n_2} z_k m_k c_j' + y_j \sum_{\ell=1}^{n_2} z_\ell c_i m_\ell' R_1' + \ell_i c_j' + c_i \ell_j' + c_i c_j'] \Bigg\} \quad i,j = 1,\ldots,n_1 \tag{B4}$$

$$\left. \begin{aligned} m_{it} m_{jt}' = m_i m_j' - \frac{1}{t-1} [2 m_i m_j' + z_j \sum_{\ell=1}^{n_1} y_\ell m_i \ell_\ell' R_2' + z_i R_2 \sum_{k=1}^{n_1} y_k \ell_k m_j' + \\ + m_i d_j' + d_i m_j'] + \frac{1}{(t-1)^2} [m_i m_j' + z_j \sum_{\ell=1}^{n_1} y_\ell m_i \ell_\ell' R_2' + \\ + z_i R_2 \sum_{k=1}^{n_1} y_k \ell_k m_j' + z_i z_j R_2 \sum_{k,\ell=1}^{n_1} y_k y_\ell \ell_k \ell_\ell' R_2' + \\ + z_i R_2 \sum_{k=1}^{n_1} y_k \ell_k d_j' + z_j \sum_{\ell=1}^{n_1} y_\ell d_i \ell_\ell' R_2' + m_i d_j' + d_i m_j' + d_i d_j'] \\ i,j = 1,\ldots,n_2 \end{aligned} \right\} \tag{B5}$$

$$\left. \begin{aligned} \ell_{it} m_{jt}' = \ell_i m_j' - \frac{1}{t-1} [2 \ell_i m_j + z_j \sum_{\ell=1}^{n_1} y_\ell \ell_i \ell_\ell' R_2' + y_i R_1 \sum_{k=1}^{n_2} z_k m_k m_j' + \\ + \ell_i d_j' + c_i m_j'] + \frac{1}{(t-1)^2} [\ell_i m_j' + z_j \sum_{\ell=1}^{n_1} y_\ell \ell_i \ell_\ell' R_2' + \\ + y_i R_1 \sum_{k=1}^{n_2} z_k m_k m_j' + y_i z_j R_1 \sum_{k=1}^{n_2} \sum_{\ell=1}^{n_1} z_k y_\ell m_k \ell_\ell' R_2' + \\ + y_i R_1 \sum_{k=1}^{n_2} z_k m_k d_j' + z_j \sum_{\ell=1}^{n_2} y_\ell c_i \ell_\ell' R_2' + \ell_i d_j' + c_i m_j' + c_i d_j'] \\ i = 1,\ldots,n_1, \quad j = 1,\ldots,n_2 \end{aligned} \right\} \tag{B6}$$

Let Λ_{ij}^t, M_{ij}^t, K_{ij}^t be defined as in (34), let c_i, $d_i = 0$ for simplicity and w.l.o.g.. We take expectations in (B4)-(B6) and drop for convenience the superscript t-1 from Λ_{ij}^{t-1}, M_{ij}^{t-1}, K_{ij}^{t-1} in the right hand side. (When taking

expectations, we use the fact that ℓ_i^{t-1}, m_i^{t-1} are independent of $y_{1,t-1}$, $y_{2,t-1}$.) We obtain:

$$\left.\begin{aligned}
\Lambda_{ij}^t &= \Lambda_{ij} - \frac{1}{t-1}\,[2\Lambda_{ij} + \sigma_j K_{ij} R_1' + \sigma_i R_1 (K_{ij})'] \\
&\quad + \frac{1}{(t-1)^2}\,[\Lambda_{ij} + \sigma_j K_{ij} R_1' + \sigma_i R_1 (K_{ij})' + \\
&\quad + \left\{\begin{array}{l} \sigma_i \sigma_j R_1 (M_{ij} + M_{ji}) R_1', \text{ if } i \neq j \\ R_1 \big(\sum\limits_{\substack{k=1 \\ k \neq i}}^{n_2} M_{kk} + E(y_i^2 z_i^2) M_{ii} \big) R_1', \text{ if } i \neq j \end{array}\right\}]
\end{aligned}\right\} \tag{B7}$$

$$i,j = 1,\ldots,n_1$$

$$\left.\begin{aligned}
M_{ij}^t &= M_{ij} - \frac{1}{t-1}\,[2M_{ij} + \sigma_j K_{ij} R_2' + \sigma_i R_1 (K_{ij})'] \\
&\quad + \frac{1}{(t-1)^2}\,\mathcal{L}_2(\Lambda_{ij}\text{'s}) \\
&\quad i,j = 1,\ldots,n_2 \text{ and } \sigma_i = 0 \text{ if } i > n_1 \\
&\qquad\qquad\qquad\qquad \sigma_j = 0 \text{ if } j > n_1
\end{aligned}\right\} \tag{B8}$$

$$\left.\begin{aligned}
K_{ij}^t &= K_{ij} - \frac{1}{t-1}\,[2K_{ij} + \sigma_j \Lambda_{ij} R_2' + \sigma_i R_1 M_{ij}] \\
&\quad + \frac{1}{(t-1)^2}\,\mathcal{L}_3(\Lambda_{ij}, M_{ij} K_{ij}\text{'s}) \; . \\
&\quad i = 1,\ldots,n_1, \quad \sigma_i = 0 \text{ if } i > n_1 \\
&\quad j = 1,\ldots,n_2, \quad \sigma_j = 0 \text{ if } j > n_2
\end{aligned}\right\} \tag{B9}$$

Defining N_t and Q as in (35), (36) we see that (B7)-(B9) can be written in compact form as in (37).

APPENDIX C

Let $x_t = (\ell'_{it},\ldots,\ell'_{n_1,t},m'_{1,t},\ldots,m'_{n_2,t})'$. Using (27) or the equivalent (B2), (B3) we have

$$x_{t+1} = x_t - \frac{1}{t}\,[R(y_{1t},y_{2t})x_t] \tag{C1}$$

where the definition of $R(y_{1t},y_{2t}) = \bar{R}_t$ is obvious from (B2), (B3). From (C1) we obtain

$$\|x_{t+1}\|^2 = \|x_t\|^2 - \frac{2}{t}\,x_t'\bar{R}_t x_t + \frac{1}{t^2}\,x_t'\bar{R}_t'\bar{R}_t x_t \tag{C2}$$

It holds

$$E\,[\|x_{t+1}\|^2 - \|x_t\|^2\,[\|x_1\|^2,\ldots,\|x_t\|^2] = \tag{C3}$$

$$= E\,[\,E\,[\|x_{t+1}\|^2 - \|x_t\|^2|\|x_1\|^2,\ldots,\|x_t\|^2]|x_1,\ldots,x_t]$$

$$E\,[x_t'\bar{R}_t x_t|\|x_1\|^2,\ldots,\|x_t\|^2] =$$

$$= E\,[\,E\,[x_t'\bar{R}_t x_t|\|x_1\|^2,\ldots,\|x_t\|^2]|x_1,\ldots,x_t] = \tag{C4}$$

$$= E\,[\,E\,[x_t'\bar{R}_t x_t|x_1,\ldots,x_t]|x_1,\ldots,x_t]|\|x_1\|^2,\ldots,\|x_t\|^2]$$

$$= E\,[x_t'E[\bar{R}_t|x_1,\ldots,x_t]x_t|\|x_1\|^2,\ldots,\|x_t\|^2]$$

$$= E\,[x_t'R_1 x_t|\|x_1\|^2,\ldots,\|x_t\|^2]$$

Since R_t depends only on y_{1t}, y_{2t} which are independent of $x_1,\dots,x_t$ and where R_1 is a constant matrix defined by

$$E\,[R(y_{1t},\, y_{2t})] = R_1 \tag{C4}$$

Similarly

$$E\,[x_t' R_t R_t x_t \,|\, \|x_1\|^2,\dots,\|x_t\|^2] = \tag{C4}$$

$$= E\,[x_t' R_2 x_t \,|\, \|x_1\|^2,\dots,\|x_t\|^2]$$

where R_2 is a constant matrix defined by

$$E\,[R'(y_{1,t},\, y_{2,t})R(y_{1t},\, y_{2t})] = R_2 \tag{C5}$$

From (C3)-(C5) we obtain:

$$E\,[\|x_{t+1}\|^2 - \|x_t\|^2 \,|\, \|x_1\|^2,\dots,\|x_t\|^2] = \tag{C6}$$

$$= E\,[x_t'(-\frac{2}{t} R_1 + \frac{1}{t^2} R_2)x_t \,|\, \|x_1\|^2,\dots,\|x_t\|^2]$$

It holds

$$-\frac{\alpha}{t} I \le -\frac{2}{t} R_1 + \frac{1}{t^2} R_2 \le \frac{\alpha}{t} I \tag{C7}$$

for some positive constant α and thus

$$|[E[x_t'(-\frac{2}{t}R_1+\frac{1}{t^2}R_2)x_t\,|\,\|x_1\|^2,\ldots,\|x_t\|^2]| \tag{C8}$$

$$\leq \frac{\alpha}{t}E[\|x_t\|^2\,|\,\|x_1\|^2,\ldots,\|x_t\|^2] = \frac{\alpha}{t}\|x_t\|^2$$

Let $V_t = \|x_t\|^2$; then from (C6) and (C7) we obtain

$$|E[V_{t+1}-V_t\,|\,V_1,\ldots,V_t]| \leq \frac{\alpha}{t} \tag{C9}$$

REFERENCES

[1] T. Basar and G. J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, 1982.

[2] T. Basar, "Equilibrium solutions in two-person quadratic decision problems with statis information structures," IEEE Trans. on Automat. Control, Vol. AC-20, 1975, pp. 320-238.

[3] G. P. Papavassilopoulos, "Solution of some stochastic quadratic Nash and Leader-Follower games," SIAM J. on Optimization and Control, Vol. 19, No. 5, Sept. 1981, pp. 651-666.

[4] I. D. Landau, "A survey of modal reference adaptive techniques-theory and applications," Automatica, Vol. 10, pp. 353-379, 1974.

[5] K. J. Astrom and B. Wittenmark, "On self tuning regulators," Automatica, Vol. 9, pp. 185-199, 1973.

[6] P. R. Kumar, "Optimal adaptive control of Linear Quadratic Gaussian systems," SIAM J. Control and Optimization, Vol. 21, No. 2, pp. 163-178, March 1983.

[7] Y. M. Chan, "Self-tuning methods for multiple controller systems," Ph.D. Thesis, Univ. of Illinois at Urbana-Champaign, Dept. of Electrical Engineering, 1981.

[8] T. L. Ting, J. B. Cruz, Jr., and R. A. Milito, "Adaptive incentive controls for Stackelberg games with unknown cost functionals," American Control Conference, June 1984, San Diego, CA.

[9] H. Robbins and S. Monro, "A stochastic approximation method," Ann. Math. Statist., 22, pp. 400-407, 1951.

[10] K. L Chung, "On a stochastic approximation method," Ann. Math. Statist., 25, pp. 463-483, 1954.

[11] J. R. Blum, "Multidimensional stochastic approximation methods," Ann. Math. Statist., 25, pp. 737-744, 1954.

[12] J. Sarks, "Asymptotic distribution of stochastic approximation procedures," Ann. Math. Statist., 29, pp. 373-405, 1958.

[13] L. Schmetterer, "Multidimensional stochastic approximation," in Multivariate Analysis--II, P. Krishnaiah (Ed.), Academic Press, 1969.

[14] M. T. Wasan, Stochastic Approximation, Cambridge Univ. Press, 1969.

[15] H. J. Kushner and D. S. Clark, Stochastic Approximation Methods for Constrained and Unconstrained Systems, Berlin; Springer-Verlage, 1978.

[16] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1970.

[17] R. B. Ash, Real Analysis and Probability, Academic Press, 1972.

[18] J. B. Harsangi, "Games with Incomplete Information Played by Bayesian Players," I-III, Management Science, Vol. 14, No. 3, Nov. 1967, pp. 159-182; Vol. 14, No. 5, January 1968, pp. 320-334; Vol. 14, No. 7, March 1968, pp. 486-502.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Not Classified	1b. RESTRICTIVE MARKINGS n/a
2a. SECURITY CLASSIFICATION AUTHORITY n/a	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 53-4503-1797	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 85-0624
6a. NAME OF PERFORMING ORGANIZATION Univ. of Southern California 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office Of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Dept. of EE-Systems, SAL 300 Los Angeles, CA 90089-0781	7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, D.C. 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Office of Scientific Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-82-0174
8c. ADDRESS (City, State and ZIP Code) Bolling Air Force Base Washington, DC 20332	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2304 TASK NO.: A6 WORK UNIT NO.:
11. TITLE (Include Security Classification) Multiobjective, Hierarchical Decision Problems in C^3	

12. PERSONAL AUTHOR(S) G. P. Papavassilopoulos

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
FINAL	FROM JUN 82 TO MAY 84	85 5 21	87

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

In the following we provide a description of the work supported by this grant. It has been announced in [1-3]. In [1] and [2] we considered the impact of changes of the information on the costs in Nash Games. (It is known that better information may be harmful in a situation of conflict, whereas it is always beneficial in the single objective-control-problem.) In [1] we show that in a general LQG statis or dynamic Nash game, better information to a player is beneficial to him if he already knows his opponent's information. We also show that in the static case, better information to a player is beneficial to him, if the new information is orthogonal to the previously known information of his and his opponent. In [2] we consider three different information structures for a static Nash game and study their interrelations. The issues addressed in [1] and [2] are of importance not only from the theoretical point of view of when, how and why changes in the information affect the objectives, but also because they provide guidelines for designing the decentralization of decisions and information of large control problems. In [3] we

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Marc Q. Jacobs	22b. TELEPHONE NUMBER (Include Area Code) (202)767-4940 22c. OFFICE SYMBOL NM

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

19. ABSTRACT (con't)

are concerned with a stochastic static Nash game in which the players do not know each others' cost parameters. We introduce a repetitive version of the problem so that as the time goes by the players update their actions – implicitly utilizing information about the parameters revealed by the previous actions of the opponent; i.e., we consider appropriately designed adaptive schemes for the actions of the players which in the limit converge – under some conditions – to the true solution of the problem. In the paper we study the convergence issues associated with such a scheme. The area of adaptive games is relatively new. It should be noticed that knowledge of one's own parameters is not always available, even more so for the parameters of an opponent. The existing work on adaptive control – i.e., single objective – problems constitutes an important background for adaptive games, which nonetheless introduce new challenging concepts and technical difficulties. We are currently continuing this line of research.

[1] "Performance versus Informativeness in LQG Multi-Objective Control Problems," with M. Tu, ACC 1984, San Diego (invited) (also submitted for publication).

[2] "Impact of explicit and Implicit Control Sharing on the Performance of Two-Person One Act LQG Nash Games," with M. Tu, Large Scale Systems, Vol. 7, 1984, pp. 219-226.

[3] "Adaptive Repeated Nash Games," Istanbul Workshop on Large Scale Systems, June 1984 (also submitted for publication in SIAM-JOC, currently under revision).

END

FILMED

10-85

DTIC